

**Pathways to products: Exploring the biosynthetic  
depth of *Myxococcus xanthus* by comprehensive  
secondary metabolite profiling**

Dissertation  
zur Erlangung des Grades  
des Doktors der Naturwissenschaften  
der Naturwissenschaftlich-Technischen Fakultät III  
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften  
der Universität des Saarlandes

von  
**Niña Socorro Cortina**

Saarbrücken

2013

Tag des Kolloquiums: 24. Mai 2013

Dekan: Prof. Dr. Volkhard Helms

Berichterstatter: Prof. Dr. Rolf Müller

Prof. Dr. Alexandra Kiemer

Vorsitz Prof. Dr. Claus Jacob

Akad. Mitarbeiter: Dr. Angelika Ullrich

Diese Arbeit entstand unter der Anleitung von Prof. Dr. Rolf Müller in der Fachrichtung  
8.2 Pharmazeutische Biotechnologie der Naturwissenschaftlich-Technischen Fakultät III  
der Universität des Saarlandes von Oktober 2011 bis März 2012.

Gedruckt mit Unterstützung des Deutschen Akademischen Austauschdienstes

# Acknowledgements

First and foremost I would like to express my deep gratitude to my adviser Prof. Dr. Rolf Müller for giving me the privilege of working on a challenging topic. I thank him for his support, encouragement, guidance and for ensuring that I had the tools and foundation I needed to carry out my work.

I would like to offer my great appreciation to Dr. Daniel Krug for being my co-adviser, for introducing me to the world of myxobacteria and comprehensive metabolomics, and for his valuable advice and input during the course of my studies.

My special thanks are extended to the Deutscher Akademischer Austauschdienst (DAAD) for the financial support and assistance during my PhD studies.

I extend my heartfelt gratitude to Dr. Alberto Plaza for his help with the structure elucidation of the myxoprincomides and for expanding my knowledge of natural products chemistry. I thank Ole Revermann for collaborating with me during my work with *Myxococcus xanthus*. I am particularly grateful to Thomas Hoffmann and Eva Luxenburger for their assistance with the LC-MS measurements. My special thanks to Wolfgang Kessler and Stephan Hüttel for carrying out the large scale fermentation of my strains and to Jennifer Herrman for performing the extensive bioactivity assays.

My deepest thanks to Janet Lei for copyediting my text and Dr. Dominik Pistorius for his valuable comments and discussion during the writing of this dissertation.

I would like to thank all of my colleagues in the Department of Pharmaceutical Biotechnology for the pleasant work environment. Special thanks to the people I have shared the office with for the support, the laughs, and the good times.

I would like to thank my friends all over the world for cheering me on, providing a kind ear, and for helping me get through the tough moments.

Lastly, I would like to dedicate this work to my family for their unwavering love and support throughout the years.

## Publications

**Cortina, N.S.,** Krug, D., Plaza, A., Revermann, O., Müller, R. Myxoprincomide, a natural product from *Myxococcus xanthus* discovered by a comprehensive analysis of the secondary metabolome. *Angew. Chem. Int. Edit.* **51**, 811 – 816 (2012).

**Cortina, N.S.,** Revermann, O., Krug, D., Müller, R. Identification and characterization of the althiomycin biosynthetic gene cluster in *Myxococcus xanthus* DK897. *ChemBioChem* **12**, 1411 – 1416 (2011).

## Conference Contributions

**Cortina, N.S.,** Revermann, O., Krug, D., Müller, R. 2011. Identification and analysis of the althiomycin biosynthetic gene cluster. *Frontiers in Medicinal Chemistry*, Saarbrücken, Germany (poster)

**Cortina, N.S.,** Revermann, O., Krug, D., Müller, R. 2010, Secondary metabolome mining: A way from gene to compound, International workshop: *Biology of Bacteria Producing Natural Products*, Tübingen, Germany (poster)

**Cortina, N.S.,** Revermann, O., Krug, D., Müller, R. 2010, Digging deep: Mining the myxobacterial secondary metabolome, 37th International Conference on the Biology of Myxobacteria: MyxoMeeting 2010, Otzenhausen, Germany (oral presentation)

**Cortina, N.S.,** Revermann, O., Krug, D., Müller, R. 2008, Identification and characterization of an althiomycin biosynthetic gene cluster in *Myxococcus xanthus* DK897, International workshop: *Biology of Bacteria Producing Natural Products*, Berlin, Germany (poster)

# Table of Contents

Acknowledgements .....	IV
Publications .....	V
Conference Contributions .....	V
Table of Contents .....	VI
List of Figures .....	XI
List of Tables .....	XIV
List of Abbreviations .....	XVI
Summary .....	1
Zusammenfassung .....	2
1 Introduction .....	3
1.1 Continuing relevance of natural product discovery .....	3
1.2 Biosynthetic logic of nonribosomal peptide synthetases and polyketide synthases .....	4
1.3 Pathway to product: biology meets natural product chemistry .....	7
1.3.1 The role of genome mining in natural product discovery .....	7
1.3.1.1 Genom isotopic approach .....	8
1.3.1.2 Targeted mutagenesis followed by metabolite profiling (TM-MP) .....	9
1.3.1.3 Heterologous expression of biosynthetic gene clusters .....	9
1.3.1.4 Activation of biosynthetic pathways .....	10
1.3.2 Genome sequence independent approaches .....	11
1.3.2.1 Proteomics-based approaches .....	11
1.3.2.2 Transcriptomics-based strategy .....	12
1.3.3 Section summary .....	13
1.4 Myxobacteria – proficient producers of secondary metabolites .....	14
1.5 The biosynthetic potential of <i>Myxococcus xanthus</i> .....	20
1.6 Principal Component Analysis .....	22
1.7 Outline of the study .....	23
1.8 References .....	25
1.9 Appendix I – List of myxobacterial compound families .....	31
2 Materials and Methods .....	32
2.1 Chemicals .....	32
2.2 Enzymes, kits, and markers .....	34
2.3 Buffers and stock solutions .....	35
2.4 Media .....	37
2.4.1 LB Medium for <i>E. coli</i> cultivation .....	37
2.4.2 Media for <i>M. xanthus</i> cultivation .....	37

2.5 Antibiotics .....	38
2.6 Instruments .....	39
2.7 Bacterial strains, oligonucleotides, and plasmids .....	40
2.7.1 Bacterial strains .....	40
2.7.2 Oligonucleotides .....	41
2.7.2.1 General primers used in this work .....	41
2.7.2.1 Primers used for the althiomycin project .....	41
2.7.2.2 Primers used for the inactivation of <i>M. xanthus</i> DK1622 biosynthetic gene clusters .....	42
2.7.3 Plasmids .....	47
2.7.3.1 General plasmids .....	47
2.7.3.2 Plasmids used for the althiomycin project .....	48
2.7.3.2 Plasmids used for the inactivation of <i>M. xanthus</i> DK1622 biosynthetic gene clusters .....	49
2.7.3.3 Plasmids used for the myxoprincomide project .....	50
2.8 <i>Myxococcus xanthus</i> .....	50
2.8.1 Mutants of <i>M. xanthus</i> DK1622 .....	50
2.8.1.1 Targeted gene inactivation of unassigned biosynthetic gene clusters .....	50
2.8.1.2 Myxoprincomide-related mutants .....	51
2.8.1.3 Althiomycin .....	51
2.8.2 Mutants of <i>M. xanthus</i> DK897 .....	52
2.8.3 Mutants of <i>M. xanthus</i> A2 .....	52
2.9 Cultivation and conservation of strains .....	52
2.9.1 Cultivation of <i>E. coli</i> .....	52
2.9.2 Cultivation of <i>M. xanthus</i> .....	52
2.10 Isolation of prokaryotic DNA .....	53
2.10.1 Isolation of genomic DNA .....	53
2.10.2 Isolation of plasmid DNA by alkaline lysis .....	53
2.10.3 Isolation of plasmid DNA by GeneJET™ Plasmid Miniprep Kit .....	53
2.11 Separation and purification of DNA .....	53
2.11.1 Separation of DNA by agarose gel electrophoresis .....	53
2.11.1 Extraction of DNA from agarose gels .....	54
2.12 Polymerase chain reaction .....	54
2.12.1 Standard PCR setup .....	54
2.12.2 PCR setup for genotypic verification of targeted gene inactivation .....	55
2.13 Enzymatic manipulation of DNA .....	57
2.13.1 Restriction endonucleases .....	57
2.13.2 Dephosphorylation .....	57

## Table of Contents

---

2.13.3 Standard ligation.....	57
2.13.4 TOPO TA cloning.....	57
2.13.5 pJET1.2 cloning.....	58
2.13.6 Phenol / chloroform extraction of ligation products.....	58
2.14 Transformation of bacteria.....	58
2.14.1 Electroporation of <i>E. coli</i> .....	58
2.14.2 Electroporation of <i>M. xanthus</i> strains .....	59
2.15 Genetic manipulation of <i>M. xanthus</i> .....	59
2.15.1 Targeted gene inactivation .....	59
2.15.2 Complementation of inactivated genes in <i>M. xanthus</i> .....	59
2.15.3 Tn5 promoter insertion by single crossover homologous recombination .....	60
2.15.4 Heterologous expression of <i>almE</i> in <i>M. xanthus</i> DK1622 .....	60
2.16 Analysis of secondary metabolite production of <i>M. xanthus</i> strains.....	60
2.16.1 Preparation of production cultures .....	60
2.16.2 LC-HRMS analysis.....	61
2.17 Sequence analysis of the <i>Myxococcus xanthus</i> PKS / NRPS biosynthetic gene clusters..	61
2.17.1 Annotation of the <i>M. xanthus</i> DK1622 biosynthetic gene clusters .....	61
2.17.2 Annotation of the althiomycin biosynthetic gene cluster from <i>M. xanthus</i> DK897 .....	61
2.17.3 Domain alignment and phylogenetic analysis .....	61
2.18 Comprehensive analysis of <i>M. xanthus</i> DK1622 secondary metabolome.....	62
2.18.1 Preparation of <i>M. xanthus</i> sample sets .....	62
2.18.2 Liquid chromatography – high resolution mass spectrometry (LC-HR-MS)..	62
2.18.3 Statistical treatment of data .....	62
2.19 Yield optimization and large scale fermentation of myxoprincomide .....	63
2.19.1 Determination of the best producer .....	63
2.19.2 Media optimization .....	63
2.19.3 Quantitative analysis of <i>M. xanthus</i> extracts .....	63
2.19.4 Strain optimization .....	64
2.19.3.2 Optimization of the introduction of XAD-16 resin.....	64
2.19.4 Large scale fermentation .....	64
2.19.4.1 300-L PS medium .....	65
2.19.4.2 9-L fed-batch VY/2 fermentation .....	65
2.20 Purification of myxoprincomide from the optimized producer strain <i>M. xanthus</i> A2.c506 .....	65
2.20.1 Isolation of myxoprincomide-c506 from 300-L PS medium fermentation of <i>M. xanthus</i> A2.c506 .....	65



2.20.2 Isolation of myxoprincomide variants from 9-L fed-batch VY/2 fermentation of <i>M. xanthus</i> A2.c506.....	66
2.21 Structure elucidation of myxoprincomide.....	66
2.21.1 NMR Analysis.....	66
2.21.2 Determination of the absolute stereochemistry of <i>N</i> -MeSer, Ser, Leu, OH-Val, Tyr, $\beta$ -Lys, and Ala.....	66
2.21.3 Determination of the absolute stereochemistry of the 2-oxo-beta-leucine in myxoprincomide-c506 .....	67
2.20.4 Tandem MS fragmentation .....	67
2.21 Feeding of labeled substrates .....	67
2.21.1 CTT (20 mL).....	67
2.21.2 VY/2 (50 mL) .....	68
2.22 Isolation and structure elucidation of althiomycin .....	68
2.23 Bioactivity assays .....	69
2.23.1 Althiomycin sensitivity assay on <i>M. xanthus</i> DK1622 wildtype and <i>attb_almE</i> mutant strains .....	69
2.23.2 Activity screening for myxoprincomide .....	69
2.23.2.1 Cell culture.....	69
2.23.2.2 Cytotoxic activity .....	69
2.23.2.3 ROS assay .....	70
2.23.2.4 NF $\kappa$ B inhibition.....	70
2.23.2.5 GR translocation .....	70
2.23.2.6 Bacterial cultures.....	71
2.23.2.7 Antimicrobial testing .....	71
2.23.3 Development Assay, sporulation frequency and motility rates .....	72
2.24 References .....	73
3 Myxoprincomide: a natural product from <i>Myxococcus xanthus</i> discovered by comprehensive analysis of the secondary metabolome.....	74
4 The molecular diversity of the myxoprincomides: linear peptides assembled by an aberrant NRPS/ PKS hybrid synthetase .....	121
4.1 Optimization of myxoprincomide production and identification of 15 additional structural variants.....	122
4.1.1 Optimization of myxoprincomide production.....	122
4.1.2 Identification of myxoprincomide variants .....	123
4.1.2.1 Group I-A: identification and characterization of myxoprincomide-c534, -c649, and -c651 .....	125
4.1.2.2 Group I-B : Identification and characterization of c513, c562 and c708 .....	129
4.1.2.3 Group I-C: Identification and characterization of myxoprincomide-c515, -c523a, c523b .....	134

## Table of Contents

---

4.2 Biosynthesis of the non-proteinogenic amino acid precursors L- $\beta$ -lysine and L- $\beta$ -hydroxyvaline .....	137
4.2.1 Formation of $\beta$ -lysine .....	137
4.2.2 Formation of L- $\beta$ -hydroxyvaline .....	140
4.3 Tn5 promoter insertion upstream of NRPS module 1 .....	140
4.4 Unusual features of the group I Mxp <sub>1622</sub> assembly line .....	142
4.4.1 Release of intermediates .....	144
4.4.2 Sluggish NRPS module and NRPS module stuttering .....	146
4.4.3 Substrate promiscuity of the A domains of Mxp <sub>1622</sub> module 2 and module 6 .....	147
.....	149
4.4.4 Revised and extended model for the biosynthesis of myxoprincomides .....	150
4.5 Bioactivity .....	151
4.5 Outlook .....	153
4.8 References.....	154
4.7 Supplementary Information.....	157
5 Identification and characterization of the althiomycin biosynthetic gene cluster in <i>Myxococcus xanthus</i> DK897.....	179
6 Conclusion & Outlook.....	212
6.1 Secondary metabolome mining .....	212
6.2 The Myxoprincomides from DK1622 .....	215
6.3 Althiomycin .....	219
6.4 Final word .....	221
6.5 References.....	221
7 Author's Contribution to Chapters 3 to 5 .....	213

## List of Figures

Figure 1.1	NRPS elongation cycle	5
Figure 1.2	PKS elongation cycle	6
Figure 1.3	Genetic capacity for PK, NRP, or hybrid PK/NRP biosynthesis prokaryotic secondary metabolite producers.	8
Figure 1.4	Summary of biological approaches to natural products discovery	13
Figure 1.5	Fruiting bodies of <i>Myxococcus xanthus</i> and <i>Chondromyces apiculatus</i>	14
Figure 1.6	Distribution of 85 compound families among myxobacterial genera	15
Figure 1.7	Percent of genera in myxobacterial secondary metabolite producers	16
Figure 1.8	Bioactive myxobacterial natural products	19
Figure 1.9	Representative structures of the <i>M. xanthus</i> DK1622 compound families	21
Figure 1.10	Scores and loadings plot	22
Figure 1.11	Finding molecular feature algorithm	23
Figure 2.1	Scheme of the genotypic verification of <i>M. xanthus</i> recombinant strains	56
Figure 3.1	Structure of myxoprincomide	74
Figure 3.2	The gene-to-compound approach via comprehensive secondary metabolome mining	76
Figure 3.3	Selection of better producer and optimization of production via promoter insertion	78
Figure 3.4	Domain organization of the newly assigned biosynthetic gene clusters and biosynthetic model for myxoprincomide assembly.	80
Figure 4.1	Tally of group I and II myxoprincomide producers	122
Figure 4.2	Comparison of myxoprincomide-c506 production in TS and VY/2 media	123
Figure 4.3	Myxoprincomide-c506 with the y- and b- fragment ions marked	124
Figure 4.4	Chromatographic profiles of myxoprincomide-c506, c534, c649, and c651	126

## List of Figures

Figure 4.5	Structure of myxoprincomide-c534	128
Figure 4.6	Proposed structure for myxoprincomide-c649 and myxoprincomide-c651	129
Figure 4.7	Chromatographic profiles of myxoprincomide-c506, -c513, c562, and c708	129
Figure 4.8	Proposed structure for myxoprincomide-c513	130
Figure 4.9	Observed $b_8$ , $y_9$ and $b_9$ product ions of the unlabeled and $d_3$ -L-leucine-labeled myxoprincomide-c562.	131
Figure 4.10	Proposed structure for myxoprincomide-c562	131
Figure 4.11	Structure of myxoprincomide-c708	134
Figure 4.12	Chromatographic profiles of myxoprincomide-c506, -c515, -c523a, and -c523b	135
Figure 4.13	Observed $y_7$ and $y_8$ product ions from the fragmentation of unlabeled and L-methionine-( $^{13}\text{C}$ -methyl)-labeled myxoprincomide-c515	135
Figure 4.14	Proposed structure for myxoprincomide-c515	136
Figure 4.15	Proposed structure for myxoprincomide-c717	136
Figure 4.16	Observed $y_7$ and $y_8$ product ions from the fragmentation of unlabeled (A) and L-methionine-( $^{13}\text{C}$ -methyl)-labeled myxoprincomide-c523a	136
Figure 4.17	Proposed structure for myxoprincomide-c523a	137
Figure 4.18	Proposed structure for myxoprincomide-c523b	137
Figure 4.19	Gene organization of MXAN_4699 and MXAN_4698	139
Figure 4.20	Extracted ion chromatograms for myxoprincomide-c506 from the wildtype, 2,3-KAM knockout mutant 4699.1, the complementation mutant 4699.3, and the L- $\beta$ -lysine-fed 4699.1 mutant strain.	139
Figure 4.21	Extracted ion chromatograms of the group I-A myxoprincomides from the wildtype and 2,3-KAM knockout mutant 4699.1.	139
Figure 4.22	Extracted ion chromatograms of myxoprincomide from the wildtype and Tn5 promoter insertion mutants A1.1, C1A.1, and C1B.1	141
Figure 4.23	Members of the group I myxoprincomides	143
Figure 4.24	Genetic environment of <i>mvp</i> <sub>1622</sub>	146

---

Figure	4.25	Biosynthetic model for the assembly of the myxoprincomides	149
Figure	4.26	Antimicrobial assay of myxoprincomide-c708 at 200 $\mu\text{g mL}^{-1}$ and determination of MIC against <i>M. hiemalis</i>	151
Figure	4.27	Fruiting body formation assay on MXAN_3779-, wildtype, and Tn5_3779 strains.	152
Figure	5.1	HPLC-MS analysis of the production of althiomycin and the dehydrated althiomycin derivative	181
Figure	5.2	Organization of the althiomycin biosynthetic gene cluster ( <i>alm</i> )	182
Figure	5.3	HPLC-MS analysis of extracts from <i>M. xanthus</i> DK897 wildtype and <i>almABCDF</i> - knockout mutants.	183
Figure	5.4	Targeted gene inactivation of <i>almD</i> (Tn5stopalmD) and <i>almC</i> (Tn5stopalmC) and LC-MS analysis.	185
Figure	5.5	Model for the biosynthesis of althiomycin in <i>M. xanthus</i> DK897	187
Figure	5.6	Althiomycin activity assay	188

## List of Tables

Table	1.1	Selected features of myxobacterial genome	18
Table	2.1	Chemicals used and their sources	32
Table	2.2	Enzymes, kits, and markers and their manufacturers	34
Table	2.3	Buffers and solutions for molecular biology	35
Table	2.4	Buffers and solutions for media preparation	36
Table	2.5	Components of LB medium	37
Table	2.6	Media used for the cultivation of <i>M. xanthus</i>	37
Table	2.7	Antibiotics used in this work	38
Table	2.8	Instruments used in this work	39
Table	2.9	Bacterial strains used in this work	40
Table	2.10	General primers used for the genotypic verification of mutants	41
Table	2.11	Primers used in the althiomycin project	41
Table	2.12	Primer list for the amplification of the ~600 bp fragment of the target gene (PKS /NRPS encoding genes)	42
Table	2.13	Control primer list for the genotypic verification of the proper integration of the knockout construct	43
Table	2.14	Primer list for the amplification of the ~600 bp fragment of the target gene (additional genes)	45
Table	2.15	Control primer list for the genotypic verification of the proper integration of the knockout construct (additional genes)	46
Table	2.16	General plasmids used in this work	47
Table	2.17	Plasmids used in the althiomycin project	48
Table	2.18	Plasmids used for the targeted gene inactivation of DK1622 biosynthetic gene clusters	49
Table	2.19	Plasmids used for the myxoprincomide project	50
Table	2.20	Mutants generated from the targeted gene inactivation of unassigned biosynthetic gene clusters	50

Table	2.21	Mutant generated for the optimized production and biosynthetic studies of myxoprincomide	51
Table	2.22	Mutant of DK1622 generated for the althiomycin project	51
Table	2.23	Mutants of <i>M. xanthus</i> DK897 generated in this work	52
Table	2.24	Mutants of <i>M. xanthus</i> A2 generated in this work	52
Table	2.25	PCR Protocols	55
Table	2.26	Reaction mixture for the genotypic verification of recombinant <i>M. xanthus</i> strains	56
Table	2.27	List of strains used for the bioactivity assays	71
Table	4.1	List of Group I myxoprincomide variants	125
Table	4.2	NMR spectroscopic data for myxoprincomide-c534 (DMSO- $d_6$ )	127
Table	4.3	NMR spectroscopic data for myxoprincomide-c708 (DMSO- $d_6$ )	132
Table	4.4	Percent identity / similarity of putative <i>M. xanthus</i> DK1622 2,3-KAM encoding genes to VioP	138
Table	4.5	Average swarming speeds (mm/day) of the <i>M. xanthus</i> DK1622 wildtype, knockout mutant, and overproducer strains	152
Table	5.1	Proteins encoded by the althiomycin biosynthetic gene cluster in <i>M. xanthus</i> DK897	182
Table	5.2	Substrate specificity of AlmA domains	183

## List of Abbreviations

<b>A domain</b>	Adenylation domain
<b>aa</b>	Amino acid
<b>ACN</b>	Acetonitrile
<b>ACP</b>	Acyl carrier protein
<b>AT domain</b>	Acyl transferase domain
<b>bp</b>	Base pairs
<b>BPC</b>	Base peak chromatogram
<b>C domain</b>	Condensation domain
<b>CID</b>	Collision induced dissociation
<b>CoA</b>	Coenzyme A
<b>COSY</b>	Correlation spectroscopy
<b>Da</b>	Dalton
<b>DAD</b>	Diode array detector
<b>D-FDLA</b>	1- fluoro-2,4-dinitrophenyl-5-D-leucinamide
<b>DMSO</b>	Dimethyl sulfoxide
<b>DNA</b>	Deoxyribonucleic acid
<b>dNTP</b>	Deoxynucleotide triphosphates
<b>DST</b>	Desalted
<b>E domain</b>	Epimerization domain
<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>EIC</b>	Extracted ion chromatogram
<b>ESI</b>	Electrospray ionization
<b>eV</b>	Electron volt
<b>h</b>	Hour
<b>HC domain</b>	Heterocyclization domain
<b>HMBC</b>	Heteronuclear multiple bond correlation
<b>HOHAHA</b>	Homonuclear Hartmann Hahn
<b>HPLC</b>	High performance liquid chromatography
<b>HR-MS</b>	High resolution mass spectrometry
<b>HSQC</b>	Heteronuclear single quantum correlation
<b>Kan</b>	Kanamycin
<b>kb</b>	Kilo base pairs
<b>kDa</b>	Kilo Dalton
<b>KS domain</b>	Ketosynthase
<b>kV</b>	Kilovolt
<b>L</b>	Liter
<b>LB</b>	Luria Bertani
<b>LC</b>	Liquid chromatography
<b>L-FDLA</b>	1- fluoro-2,4-dinitrophenyl-5-L-leucinamide
<b>M</b>	Molar
<b>Mbp</b>	Mega base pairs
<b>MDa</b>	Mega Dalton
<b>mDa</b>	Milli Dalton
<b>MeOH</b>	Methanol
<b>MF</b>	Molecular formula
<b>MIC</b>	Minimum inhibitory concentration
<b>min</b>	Minute



<b>mM</b>	Millimolar
<b>MS</b>	Mass spectrometry
<b>MS/MS or MS<sup>2</sup></b>	Tandem mass spectrometry
<b>MT</b>	Methyltransferase
<b>MW</b>	Molecular weight
<b><i>m/z</i></b>	Mass-to-charge
<b>NMR</b>	Nuclear magnetic resonance
<b>NRPS</b>	Nonribosomal peptide synthetase
<b>OD</b>	Optical density
<b>Ox domain</b>	Oxidation domain
<b>PCA</b>	Principal component analysis
<b>PCP</b>	Peptidyl-carrier protein
<b>PCR</b>	Polymerase chain reaction
<b>PKS</b>	Polyketide synthase
<b>Ppant</b>	4'-phosphopantetheinyl
<b>ppm</b>	Parts per million
<b>PST</b>	Peptide sequence tag
<b>RNA</b>	Ribonucleic acid
<b>ROESY</b>	Rotating frame Overhauser effect spectroscopy
<b>rpm</b>	Revolutions per minute
<b>RT</b>	Retention time
<b>SAM</b>	S-adenosylmethionine
<b>SAR</b>	Structure-activity relationship
<b>sec</b>	Second
<b>TE domain</b>	Thioesterase domain
<b>TFA</b>	Trifluoroacetic acid
<b>TOCSY</b>	Total correlation spectroscopy
<b>TM-MP</b>	Targeted mutagenesis – metabolite profiling
<b>UPLC</b>	Ultra high performance liquid chromatography
<b>UV</b>	Ultraviolet
<b>V</b>	Volt
<b>v/v</b>	Volume per volume
<b>WT</b>	Wildtype
<b>Zeo</b>	Zeocin
<b>Å</b>	Angstrom
<b>°C</b>	Degrees Celsius
<b>2,3-LAM</b>	2,3-lysine aminomutase
<b>μF</b>	Microfarad
<b>μL</b>	Microliter
<b>μM</b>	Micromolar
<b>Ω</b>	Ohm



## Summary

Microbial secondary metabolites are a rich source of bioactive compounds which act on novel therapeutic targets. The availability of whole genome sequences revealed that, thus far, the expanse of the secondary metabolite biosynthetic potential of bacteria has been grossly underestimated. In this work, the biosynthetic potential of the myxobacterial species *Myxococcus xanthus* was investigated.

Examination of the *M. xanthus* DK1622 genome revealed 18 biosynthetic gene clusters, only 5 of which have been previously assigned to compound classes. In the present work the application of a comprehensive secondary metabolite profiling approach in combination with molecular biological methods facilitated the association of 3 novel *M. xanthus* DK1622 compounds to their corresponding biosynthetic pathways. Complete structure elucidation was carried out for one selected candidate and revealed an unusual peptide, named myxoprincomide-c506. A biosynthetic model for myxoprincomide was proposed based on the giant multidomain NRPS/PKS enzyme Mxp from *M. xanthus* DK1622. Further investigations established myxoprincomide-c506 as the founding member of an amazingly diverse metabolite family, and 15 additional myxoprincomide variants were identified. Their formation comprises both rare biosynthetic building blocks and intriguing deviations from textbook biosynthetic logic.

Furthermore, retrobiosynthetic analysis and genome mining of the *M. xanthus* DK897 genome led to the first identification of an althiomycin biosynthetic pathway encoded by the *alm* gene cluster. Following confirmation by targeted gene inactivation coupled to metabolite profiling of the mutants, a biosynthetic model for althiomycin assembly by a hybrid NRPS/PKS system was proposed.

## Zusammenfassung

Sekundärmetabolite mikrobiellen Ursprungs sind eine ergiebige Quelle bioaktiver Verbindungen, die neuartige therapeutische Angriffspunkte enthüllen. Die Verfügbarkeit kompletter Genomsequenzen hat offenbart, dass das Potential von Bakterien zur Biosynthese von Sekundärmetaboliten bis dato bei weitem unterschätzt worden ist. In dieser Arbeit ist das biosynthetische Potential der myxobakteriellen Gattung *Myxococcus xanthus* eingehend untersucht worden.

Die Auswertung des Genoms von *M. xanthus* DK1622 lässt 18 Biosynthesegencluster erkennen, aber nur fünf dieser Gencluster konnten bisher Verbindungsklassen zugeordnet werden. Die Anwendung eines umfassenden Ansatzes zur Analyse von Sekundärstoffprofilen in Kombination mit molekularbiologischen Methoden, ermöglichte in dieser Arbeit die Zuordnung von drei neuen Verbindungen von *M. xanthus* DK1622 zu den entsprechenden Biosynthesewegen. Die Verbindung Myxoprincomid-c506 wurde für eine komplette Strukturaufklärung ausgewählt und wird hier als neuartiges Peptid mit mehreren seltenen Strukturmotiven beschrieben. Ein Modell der Myxoprincomid-Biosynthese in *M. xanthus* DK1622 wurde basierend auf dem 1.6 MDa großen NRPS/PKS Multidomänenenzym Mxp aufgestellt. Es konnten außerdem 15 weitere Myxoprincomid-Derivate identifiziert werden. Die außergewöhnliche strukturelle Diversität dieser neuen Substanzfamilie beruht dabei auf seltenen Biosynthese-Bausteine und ungewöhnlichen Biosyntheseschritten.

Außerdem führten retrobiosynthetische Überlegungen und die Analyse des Genoms von *M. xanthus* DK897 zur Identifizierung des Althiomycin Biosyntheseweges, der vom *alm* Gencluster kodiert wird. Nach gezielter Inaktivierung des Genclusters und Analyse der Mutanten bezüglich ihrer Althiomycinproduktion, konnte ein Modell für die NRPS/PKS-basierte Biosynthese von Althiomycin aufgestellt werden.

# 1 Introduction

## 1.1 Continuing relevance of natural product discovery

Since the ancient times, the use and efficacy of natural products (NPs) for human therapeutics has been unrivaled.<sup>[1]</sup> Nevertheless, for the past three decades, the pharmaceutical industry deemed natural products-based drug discovery as a conquered horizon and instead refocused its drug discovery programs in combinatorial chemistry and high throughput screening.<sup>[2,3]</sup> Without discounting the successes of synthetic drugs, natural products made a more significant impact in therapeutics during this period. From 1981-2006, the majority of the new chemical entities (NCEs) approved for drug use are natural products, their derivatives, and NP-inspired synthetic structures obtained from plant, fungal, and microbial sources.<sup>[4]</sup>

Most notable is the high percentage of antibacterial and anticancer agents among the microbial derived bioactive secondary metabolites (SMs), some of which are even promising drug leads. <sup>[5,6]</sup> Majority of these bioactive compounds fall under 2 natural product classes: nonribosomal peptides (NRPs), and polyketides (PKs). NRPs and PKs offer multifaceted binding properties due to their similarity to the binding surface of the drug target (NRPs) and their structural diversity and complex stereochemistry (PKs).<sup>[7]</sup>

NRPs and PKs are products of mechanistically-cognate thiotemplated modular megasynthases. The rapid progress in sequencing technologies and the tools for annotation of bacterial genomes hastened the process of the exploration of these biosynthetic pathways. With these advances it became obvious that the breadth and depth of bacterial biosynthetic potential was severely underestimated, not just in terms of numbers but concerning the structural diversity they can provide.<sup>[8]</sup>

In times of increasing numbers of multidrug-resistant pathogens as well as increasing incidences of cancer, due to a higher life expectancy in the industrialized countries, there is a need to reevaluate and restrategize the methods by which these diseases can be treated. The new strategies entail the identification of new targets for which new chemotypes are required.<sup>[7]</sup> The source is boundless; it is a matter of access.

One of the initial challenges for drug development from NRPs and PKs is the identification of these bioactive compounds from the complex matrices of microbial culture broths. The next challenge is to obtain sufficient amounts for preclinical

development and clinical studies, which even after large scale fermentation of the producer, is an arduous task.<sup>[7]</sup> New techniques, from improved molecular biology to sophisticated analytics, aim to address this multifaceted problem.

### 1.2 Biosynthetic logic of nonribosomal peptide synthetases and polyketide synthases

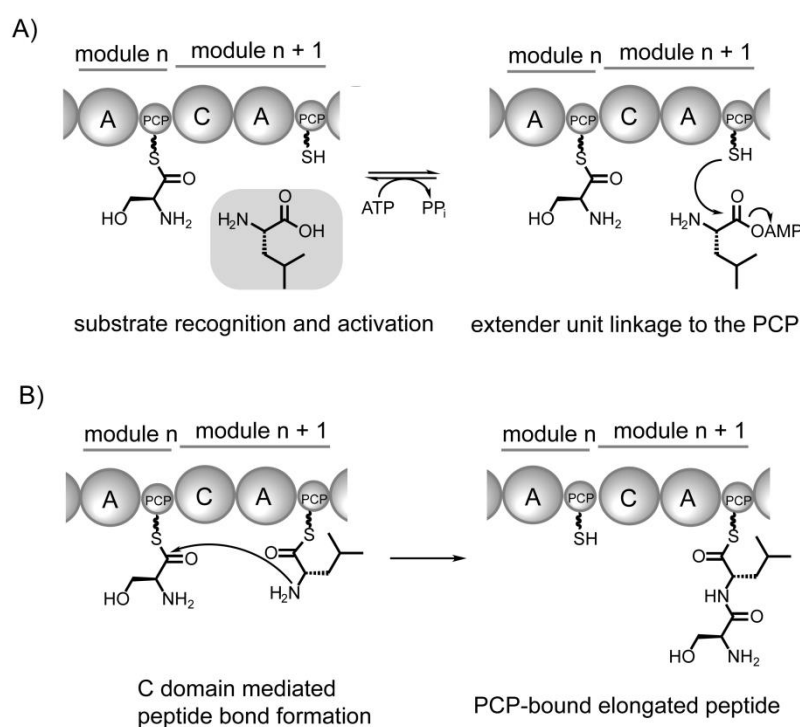
Nonribosomal peptide synthetases (NRPSs) are large multifunctional enzymes that assemble peptidic natural products. Building block monomers are not only limited to the 20 proteinogenic amino acids but can include nonproteinogenic, hydroxylated, *N*-methylated, and D-amino acids.<sup>[9]</sup>

NRPSs are subdivided into modules which are responsible for the incorporation of one amino acid into the final product. These modules are comprised of catalytically independent domains, distinguishable by highly conserved core motifs at the protein level. Each domain carries out the successive reactions necessary for peptide elongation.<sup>[10]</sup> The minimum set of domains for an elongation cycle consists of an adenylation (A) domain, a peptidyl carrier protein (PCP), and a condensation (C) domain. (Fig. 1.1)

The A domain plays a central role in determining the primary sequence of the nonribosomal peptide by selecting and activating the substrate as an aminoacyl adenylate. The 8-10 amino acids most accessible on the protein surface to the amino acid monomer, referred to as the specificity conferring code, dictates the cognate amino acid to be incorporated into the nascent chain.<sup>[11;12]</sup> NRPS annotation tools such as PKS/NRPS Analysis and NRPSPredictor2 (and its predecessor NRPSPredictor) use this nonribosomal code to provide *in silico* substrate predictions for the A domains.<sup>[13-15]</sup>

The activated aminoacyl substrate is transferred and covalently bound to the thiol group of the 4'-phosphopantetheinyl (Ppant) arm of the PCP. The PCP enables the transport of the activated amino acids between the catalytic centers of the NRPS. The PCP is posttranslational modified to its *holo* form via the transfer of the Ppant cofactor to the catalytic serine residue of the *apo*-PCP, a reaction that is catalyzed by a dedicated 4'-phosphopantetheinyl transferase.<sup>[16;17]</sup>

Elongation is accomplished by the C domain wherein the free amino group of the downstream monomer attacks the carbonyl-thioester of the upstream substrate to form a peptide bond. Alternatively, the dual process of condensation and cyclization to form oxazoline and thiazoline moieties from serine, threonine, and cysteine residues can be carried out by heterocyclization (HC) domains in a mechanism that is yet to be elucidated.<sup>[18]</sup> Formation of an oxazole or thiazole requires an oxidation domain typically embedded in an A domain. Formation of thiazoles as a post translational modification is demonstrated in ribosomally synthesized peptides as an ATP-dependent cyclodehydration reaction.<sup>[19]</sup>



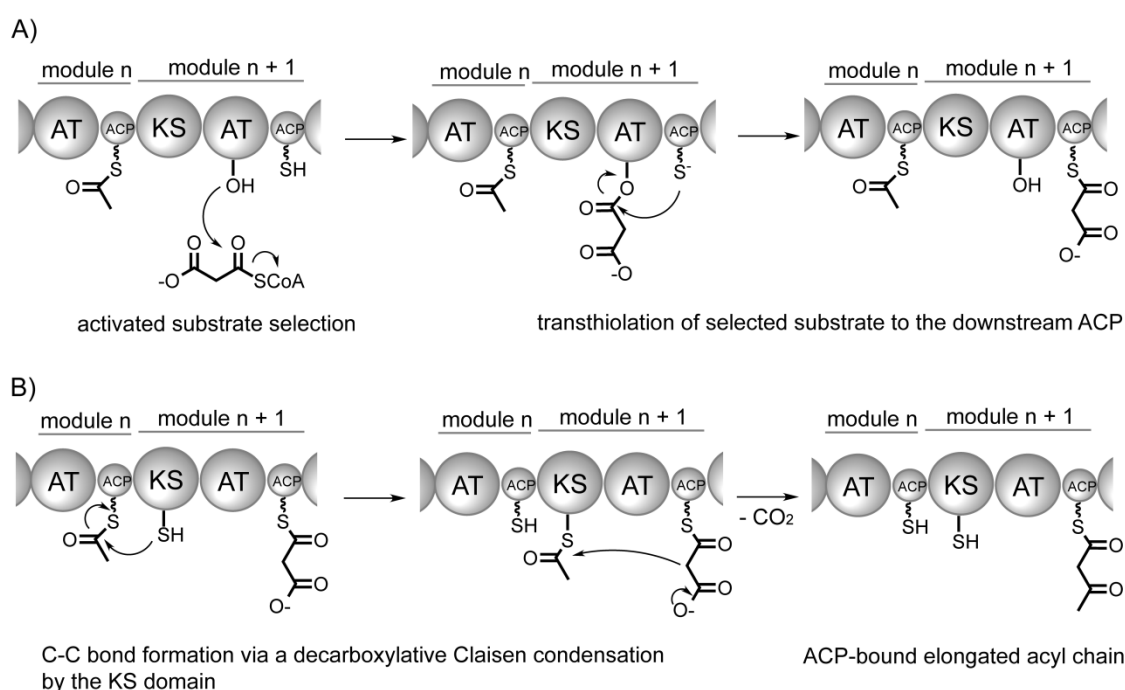
**Figure 1.1:** NRPS elongation cycle. A) Amino acid recognition, activation and loading. B) Peptide bond formation and elongation.

The structural diversity of NRPs can be further expanded by modifications of the elongating peptide carried out by methyltransferases (MT), oxidation (Ox), and epimerization (E) domains. Release of the final product is performed by the terminal thioesterase (TE) domain either by hydrolysis or by macrocyclization.<sup>[20]</sup>

Microbial polyketides are derived from acyl-CoA precursors (e.g. malonyl-CoA or methylmalonyl-CoA) that undergo sequential decarboxylative condensations assembled by modular polyketide synthases (PKSs) in a manner similar to fatty acid biosynthesis.<sup>[21]</sup> One elongation cycle requires an acyl transferase (AT) domain to select the starter or

## 1 Introduction

extender units and loads them onto the Ppant arm of the acyl carrier protein (ACP). These ACP-tethered acyl moieties undergo C-C bond formation via a decarboxylative Claisen-type condensation accomplished by the ketosynthase (KS) domain.<sup>[22]</sup> (Fig. 1.2) Akin to NRPS A domains, substrates for AT domains can be predicted *in silico* by sequence analysis.<sup>[23-26]</sup> Auxiliary domains like  $\beta$ -ketoacyl reductase (KR),  $\beta$ -hydroxyacyl dehydratase (DH), and enoyl reductase (ER) carry out the reductive processing at the  $\beta$ -carbonyl centers. The polyketide skeleton is released by TE domains as linear structures or as macrolides. Post assembly modifications such as methylations, hydroxylations, and glycosylation can be crucial for the full biological activity to the final product.<sup>[27]</sup>



**Figure 1.2:** PKS elongation cycle. A) Selection and loading of an an activated extender unit. B) Elongation of the polyketide chain by a Claisen-type condensation of the extender unit

To date, PKS systems have been categorized into three types based on their products and assembly mechanisms. Type I PKSs are giant multifunctional enzymes which generate macrolides, polyethers, and polyenes either in a linear or iterative manner akin to the mechanism presented in Figure 1.2.<sup>[22]</sup> Type II PKS systems use multienzyme complexes composed of a set of discrete  $KS_{\alpha}$ / $KS_{\beta}$ /ACP proteins which act iteratively to catalyze the formation of poly- $\beta$ -keto intermediates and subsequent cyclization results in aromatic moieties.<sup>[28]</sup> The chalcone synthase-like Type III PKSs are composed of homodimeric KS domains with a single active site that performs the iterative loading and decarboxylative condensation of free acyl-CoA thioesters.<sup>[29]</sup>



The structural diversity of PKS- and NRPS-based scaffolds can be further expanded by the combination of these two compatible megaenzymes into hybrid NRPS/PKS systems (heretofore referred to as hybrid systems). Extension steps require catalytic leniency where the KS and C domains are able to perform the condensation between acyl and amino acid substrates at the hybrid interface.<sup>[22]</sup>

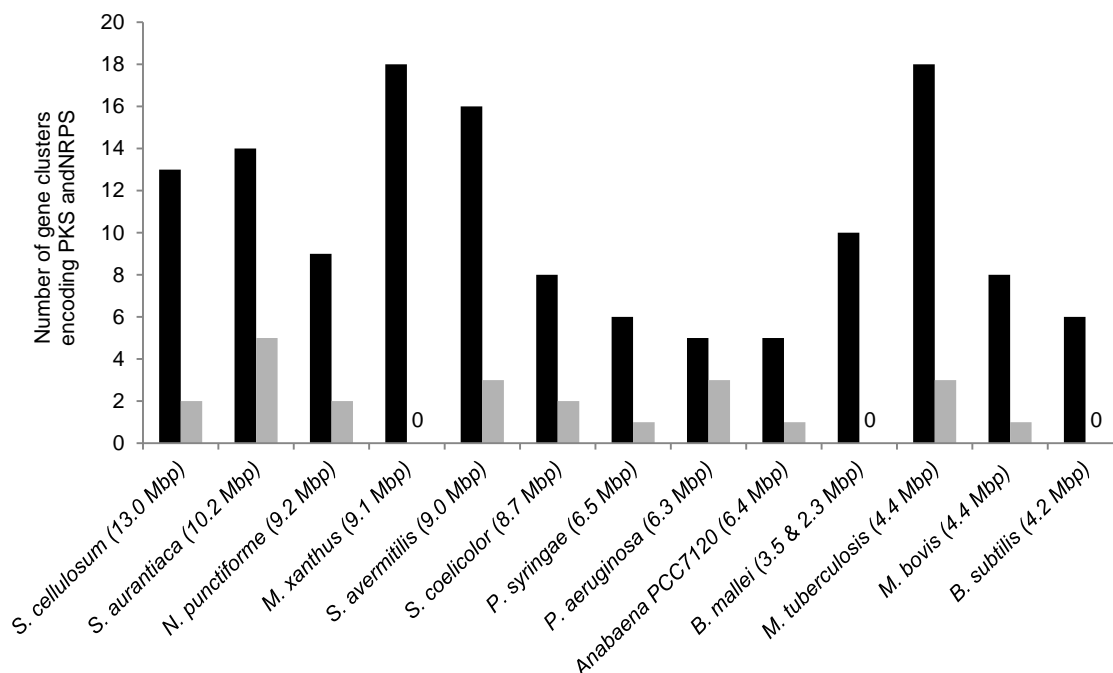
### 1.3 Pathway to product: biology meets natural product chemistry

The classical strategy to microbial natural product discovery is bioassay-guided serial fractionation of extracts. Although this is a successful approach in discovering bioactive compounds, the rate of rediscovery of known compounds is nowadays quite high.<sup>[30]</sup> Dereplication, the rapid identification of known compounds, remains a constant analytical challenge.<sup>[31,32]</sup> Fortunately, during the past two decades, biological tools were developed to complement analytical chemistry in natural products discovery. To provide a small glimpse into the available tools, a few of the available technologies are highlighted in this section.

#### 1.3.1 The role of genome mining in natural product discovery

The discovery of the three genes encoding the biosynthetic pathway for the PK-derived antibiotic erythromycin (6-deoxyerythronolide B synthase or DEBS) from *Saccharopolyspora erythraea* over 20 years ago was a milestone in the understanding of assembly line enzymology. Parallel discoveries revealed that NRPs are constructed via an analogous modular mechanism.<sup>[8]</sup>

The genes encoding these assembly lines are typically clustered in a locus in the prokaryotic genome. The role of molecular biology in NP discovery has since expanded, providing a new gateway with genome mining for PKS and NRPS pathways serving as the central tool.<sup>[33-37]</sup> Genome annotation also revealed that the number of encoded biosynthetic pathways is significantly greater than the number of known compounds from producer strains.<sup>[8,30]</sup> (Fig. 1.3)



**Figure 1.3:** Genetic capacity for PK, NRP, or hybrid PK/NRP biosynthesis of selected fully sequenced prokaryotic secondary metabolite producers. Identified PKS and/or NRPS biosynthetic gene clusters are shown in black bars. Known PK, NRP, or hybrid PK/NRPS compounds from the sequenced strains at the time of genome sequencing are in grey.<sup>[35]</sup>

In an ideal setting *in silico* prediction of the associated product could be performed if the biosynthetic pathway follows the colinearity rule where the number and type of monomers as well as the expected modifications can be inferred from the domains and tailoring domains identified.<sup>[30]</sup> For the most part this is not the case and several methods have been devised to improve the correlation between pathways and products.

### 1.3.1.1 Genomisotopic approach

The genomisotopic approach combines genome sequence analysis and feeding of isotope-labeled precursors followed by NMR-guided fractionation to determine the products of orphan biosynthetic gene clusters. The strategy led to the correlation of the antibiotic orfamide A to its biosynthetic gene cluster *ofa* in *Pseudomonas fluorescens* Pf-5.<sup>[38]</sup> The method relies heavily on *in silico* A domain prediction of cognate amino acids to select the isotope labeled substrates. This dependence highlights one of the method's limitations: not all A domain substrates can be predicted *in silico*. In addition, it is also likely that a labeled amino acid serves as a precursor for several orphan pathways and therefore direct association of gene to product could be hampered. Last but not least it is strictly restricted to NRPS pathways.

### 1.3.1.2 Targeted mutagenesis followed by metabolite profiling (TM-MP)

Targeted gene inactivation followed by comparative metabolite profiling is one of the “classical” methods in correlating secondary metabolites to their gene clusters and vice versa. The method is particularly applicable to biosynthetic pathways which deviate from the colinearity rule or whose domains could not be fully characterized. A gene cluster is inactivated by insertional or deletion mutagenesis and the phenotype is characterized by comparing the metabolite profiles of the wildtype and recombinant strains by high performance liquid chromatography – mass spectrometry (HPLC-MS).<sup>[30]</sup> The sequence information of the cluster does not necessarily have to come from whole genome sequences. For example, early discovered myxobacterial biosynthetic pathways were based on sequences obtained from cosmid libraries or recovered vectors post transposon mutagenesis.<sup>[39-42]</sup>

One of the very first examples of novel secondary metabolites discovered by wildtype/mutant comparative metabolite profiling is the discovery of myxochromide S from *Stigmatella aurantiaca* DW4/3-1.<sup>[43]</sup> The biosynthetic pathway was identified from a cosmid library but the corresponding product was unknown. With the sequence information at hand, targeted gene inactivation was carried out and identification of myxochromides S<sub>1-3</sub> was achieved by liquid chromatography-diode array detection-mass spectrometry (LC-DAD-MS) analysis of the wildtype and mutant strain extracts. As myxochromides S<sub>1-3</sub> represent major metabolites in the wildtype, the absence of their corresponding peaks in the chromatogram of the mutant strain was easy to detect.

This strategy can lead to rapid gene to compound correlation if the microorganism of interest is amenable to mutagenesis and the analytical method used is capable of detecting the phenotype.

### 1.3.1.3 Heterologous expression of biosynthetic gene clusters

Successful discovery of natural products by heterologous expression was achieved from relatively small gene clusters (e.g. the sesquiterpene avermitilol from *Streptomyces avermitilis* and the ribosomally-derived depsipeptide microviridins from *Microcystis* species).<sup>[44;45]</sup> PKS and NRPS gene clusters are by no means small.

Although still limited by size constraints, heterologous expression of entire PKS / NRPS biosynthetic pathways has been effective using Red/ET recombineering. Examples include the heterologous expression of the myxochromide S and myxothiazol

biosynthetic gene clusters from *Stigmatella aurantiaca* DW 4/3-1 into *Pseudomonas putida* and *Myxococcus xanthus* DZF1 respectively, which resulted in high level production of the corresponding products.<sup>[46;47]</sup>

The discovery of novel PKS / NRPS derived SMs by heterologous expression has only been recently achieved by heterologously expressing biosynthetic pathways from *Photorhabdus luminiscens*. Using a full length RecE/RecT based direct DNA cloning strategy, 10 megasynthetase gene clusters were heterologously expressed in *E. coli*, 2 of which produced the novel compounds luminmycin A and luminmide A/B.<sup>[48]</sup>

### 1.3.1.4 Activation of biosynthetic pathways

Modification of cultivation conditions allows access to secondary metabolites whose biosynthetic pathways are not expressed under standard fermentation conditions. For example, in the OSMAC (One Strain MANY Compounds) approach factors such as the media, aeration, temperature, and even the shape of the culture flask are varied to obtain a broader secondary metabolite spectrum.<sup>[49]</sup> Additional biosynthesis inducers include external stressors (e.g. enzyme inhibitors), mimicry of the natural substrate, interspecies crosstalk by co-culturing, and manipulation of regulators to switch on the biosynthetic pathways.<sup>[50]</sup>

The discovery of the prenylated quinoline-2-one alkaloids aspoquinolones from *Aspergillus nidulans* was facilitated by a combination of genome mining, growth in different culture conditions, and analytical techniques.<sup>[51]</sup> (Three genes encoding proteins with high sequence similarity to anthranilate synthase (AS) were detected. Anthranilate synthase catalyzes the conversion of chorismate to anthranilic acid, a precursor in the biosynthesis of quinazoline, quinolone, and acridine alkaloids. The presence of multiple AS-encoding genes led to the assumption that some of them were involved in secondary metabolite biosynthesis but their putative products were unknown. This prompted a HPLC--UV-MS screen of 40 *A. nidulans* HKI 0410 extracts grown under 40 different culture conditions. One growth condition (solid state fermentation on rice medium) revealed two unknown compounds which, after isolation and structural elucidation, were determined to be novel prenylated quinolone alkaloids named aspoquinolone A and B. Two additional structural variants (aspoquinolone C and D) were identified after large-scale fermentation.

Activation of a silent PKS gene cluster by regulatory cross talk led to the discovery of the novel secondary metabolite asperfuranone.<sup>[52]</sup> The production of asperfuranone was previously not detected under standard and modified media conditions. The activation of the cluster was serendipitously observed when the regulatory gene *scpR* for a different cryptic NRPS gene cluster was being investigated and two novel compounds, not related to the NRPS being investigated but appear to correlate with the gene product of the PKS gene cluster, were identified.<sup>[53]</sup>

### 1.3.2 Genome sequence independent approaches

The challenge of biosynthetic pathway discovery is not only addressed by genome dependent methods, as “protein-” and “transcript-first” strategies are being developed. The main advantage of these two approaches is that gene clusters which are actively transcribed and expressed under standard cultivation conditions are detected. It is necessary to point out however that detection of expressed pathways or transcribed genes does not necessarily equate to detection and identification of novel compounds.

#### 1.3.2.1 Proteomics-based approaches

The method called PrISM (Proteomic Investigation of Secondary Metabolism) takes advantage of the facile release of the Ppant cofactor from carrier proteins (CPs) by tandem MS to identify sequences attributed to NRPS and PKS systems.<sup>[54]</sup> Proteomes were separated by SDS-PAGE and protein bands with MW > 150 kDa were excised for in gel tryptic digest and analyzed by LC-MS/MS. MS<sup>n</sup> experiments identified peptides with Ppant elimination ions and provided sufficient information for *de novo* sequencing. PCR amplification using degenerate primers can afford the rest of the sequence information. Functional analysis of the gene cluster and tandem MS data were used to identify the products of the detected pathways. The method was successful in identifying novel secondary metabolites (koranimine and a new family of lipoheptapeptides) from *Bacillus* species.<sup>[55]</sup> An improved PrISM approach which eliminated Coomassie detection and included OMSSA (*Open Mass Spectrometry Search Algorithm*) database search and improved *de novo* sequencing facilitated the identification of the biosynthetic gene clusters for the siderophore foroxymithine and the patented antibiotic S 213 L in actinomycetes.<sup>[56]</sup>

An alternative proteomics-based approach called orthogonal active site identification system (OASIS) takes advantage of activity based proteome profiling (ABPP) to identify

PKSs and NRPSs.<sup>[57]</sup> Small molecule “warheads” utilize the endogenous catalytic domains and react irreversibly by covalent interaction with the active site residue.<sup>[58]</sup> CPs are labeled by azido-CoA precursors while TEs are targeted by fluorophosphonate probes. The labeled proteins undergo avidin enrichment and on bead tryptic digest and probe bearing peptides are identified by multidimensional protein identification technology (MudPIT) LC-MS/MS. A recent update of the OASIS technology combines ABPP approach by labeling solely the TE domains and detecting the PPant ejection ion from tandem MS analysis, similar to the PrISM method, in identifying fragments from the terminal ends of PKSs and NRPSs.<sup>[58]</sup>

PrISM and OASIS have a common limitation: both can detect expressed PKSs and NRPSs but a significant amount of additional work is still required to determine the associated products of these modular synthases. The PrISM method claims simultaneous identification of biosynthetic pathways and products. However, a close inspection of their methodology shows that the process, although shown to be successful, is not straightforward, requiring reverse translation and sequence analysis prior to product identification. OASIS has not at all made significant advances in this regard and has stopped at CP or TE detection. The OASIS method also suffers from the promiscuity of the small molecule probes, reacting with other proteins aside from the PKS or NRPS targets (i.e. FP-probes for TE react readily with other serine hydrolases).

### 1.3.2.2 Transcriptomics-based strategy

A recent study demonstrated the use of the messenger RNA (mRNA) transcript for biosynthetic pathway identification. The screening strategy utilized the conditional transcriptome, a collection of cDNA based on total RNA isolated at a particular time point during cultivation. PCR amplification using degenerate primers was performed on the selectively-enriched transcript of NRPS / PKS genes. The sequence of six amplicons could not be correlated to any known biosynthetic gene clusters. Targeted inactivation based on these sequences resulted in the same phenotype: loss of production of mycotrienin (MYC) related compounds. Using the PCR sequences as probes, the *myc* gene cluster was obtained and annotated. Deletion mutants spanning *mycA1-mycD2* confirmed the association of the *myc* cluster to the MYC products.<sup>[59]</sup>

The main advantage of this approach is that it requires no a priori knowledge of the genome sequence and could readily be applied to any secondary metabolite producing

strain. The method could easily be extended to examine the conditional transcriptome and secondary metabolite production at different time points during cultivation. Nevertheless, there is detection bias for the enriched NRPS and PKS genes transcribed and there is a risk to overlook pathways with a lower rate of transcription.

### 1.3.3 Section summary

Natural products biology has come a long way since the discovery of the DEBS pathway. Individually, the available tools have their advantages and drawbacks. (Fig. 1.4)

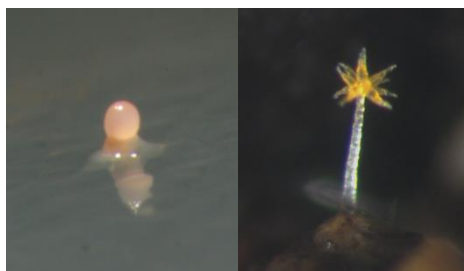
But combined, these methods contributed significantly to the immense progress in natural product discovery. A few methods mentioned in this section were instrumental in the advances in myxobacterial secondary metabolite research.



**Figure 1.4:** Summary of the biological approaches to natural product discovery mentioned in this review with their advantages (+) and drawbacks (-). Common to all methods is the tremendous task of discovering their corresponding products whether if the starting data comes from the genome, the proteome, or the transcriptome.

### 1.4 Myxobacteria – proficient producers of secondary metabolites

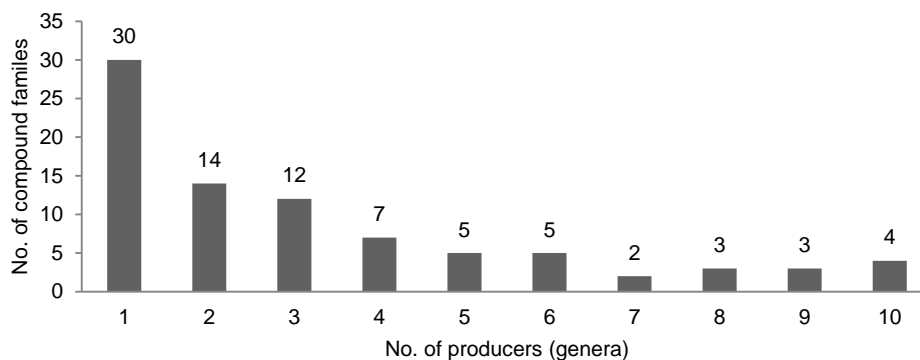
Myxobacteria are Gram-negative  $\delta$ -proteobacteria which move by gliding or creeping over surfaces forming swarms that spread into unoccupied areas. Soil and dung provide an optimal habitat for bacteriolytic and proteolytic species while cellulose degraders thrive on bark and decaying plant material.<sup>[60]</sup> Myxobacterial species have also been isolated from marine environments (i.e. members of the genus *Haliangium*) as well as from a variety of terrestrial biomes<sup>[61-64]</sup> When myxobacteria experience starvation conditions, the swarms begin to form aggregates and pile up to produce fruiting bodies whose morphologies ranges from simple knobs (e.g. genus *Myxococcus*) up to sophisticated tree-like formations (e.g. genus *Chondromyces*) (Fig. 1.5). Encased in these fruiting bodies are dessicant-resistant myxospores which guarantee the survival of the colony until nutrient levels return to normal.<sup>[65]</sup>



**Figure 1.5:** Fruiting bodies of *Myxococcus xanthus* (A) and *Chondromyces apiculatus* (B). Photos courtesy of Ronald Garcia.

Aside from their remarkable social development, myxobacteria are known for their exceptional ability to produce novel secondary metabolites.<sup>[66]</sup> Majority of these compounds are not only distinct from those produced by other microbial sources, but appear to be diversified within myxobacterial genera as well. (Fig. 1.6)

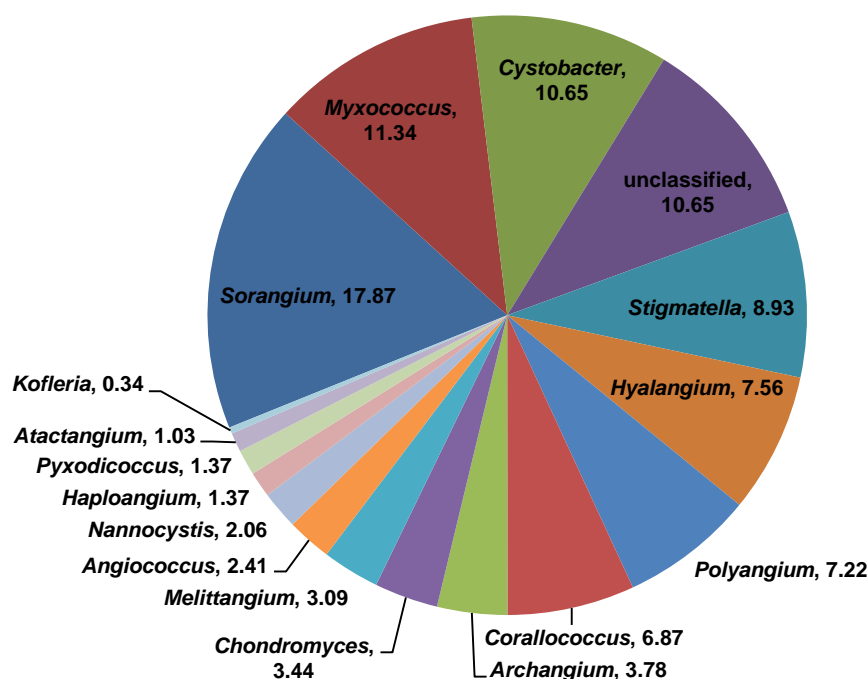




**Figure 1.6:** Distribution of 85 compound families (Appendix I) among myxobacterial genera. The y-axis counts the number of individual compound families produced by species from one or more genera (x-axis). Thirty of the compound families are uniquely produced by members of a single genus while 4 compound families are common to 10 myxobacterial genera. Data based on Myxobase (stand: 5. January 2012)

Secondary metabolites produced by the genera *Sorangium*, *Myxococcus*, *Cystobacter*, and *Stigmatella* dominate the myxobacterial compound spectrum admittedly because these genera were predominantly cultivated in the past. However the contributions of the other genera are steadily expanding. (Fig. 1.7)

In terms of numbers of compounds discovered, myxobacteria lag behind the more intensively studied secondary metabolite producers like actinomycetes, fungi, and bacilli.<sup>[67]</sup> The research pace is challenged by the slow growth of myxobacterial strains (4-14 hours doubling time) and the tedious process of re-culturing to obtain a homogeneous suspension. These two issues impact large scale fermentation which is necessary to obtain sufficient amount of pure compound for structure elucidation.<sup>[68]</sup>



**Figure 1.7:** Percent of genera in myxobacterial secondary metabolite producers. Redundancy due to structural variants has been reduced by summarizing the derivatives into one compound family. Data was based on Myxobase (stand: 5. January 2012)

Myxobacteria are microbial predators which prey on other bacteria and yeast and use them as secondary substrates.<sup>[65;67]</sup> Majority of bioactive myxobacterial natural products act against yeast/fungi (~54%) and bacteria (~29%), a finding that might be interpreted as a consequence of the competitive pressures of their natural habitat.<sup>[69]</sup> A recent study revealed that myxobacteria proactively use secondary metabolites as predatory weapons.<sup>[70]</sup>

The primary mode of action for antifungal activity is the inhibition of electron flow within the mitochondrial respiratory chain by specifically targeting NADH dehydrogenase (complex I) and/or cytochrome  $bc_1$  (complex III), a mechanism quite common to myxobacterial antifungal compounds but rarely observed with fungicides from other bacterial sources.<sup>[71]</sup> Other targets include interruption of DNA replication and transcription (leupyrrins), inhibition of translation (gephyronic acid), and disruption of the cell membrane (ambruticin).<sup>[72-74]</sup> The success story of soraphen, a polyketide which uniquely targets fungal acetyl-CoA carboxylase that was investigated as fungicidal agent in agriculture, was slashed by observed side effects like teratogenic and allergenic activity on test animals halting any further development.<sup>[66]</sup>

The antibacterial modes of action of myxobacterial natural products are just as diverse as the fungicidal ones. Thuggacins target bacterial electron transport and has been shown to be active against *Mycobacterium tuberculosis*.<sup>[75]</sup> Corallopyronin, myxopyronin, and ripostatin all interact with the switch region of bacterial RNA polymerase thereby interfering with the initiation of mRNA synthesis.<sup>[76]</sup> Protein synthesis is inhibited by althiomycin at the peptidyl transfer stage<sup>[77]</sup> whereas the potent antibiotic myxovalargin disrupts the binding of the aminoacyl-tRNA to the A site of the ribosome. In addition, myxovalargin has at higher concentration and prolonged incubation a secondary effect of damaging the cell membrane.<sup>[78]</sup>

Ten percent of myxobacterial secondary metabolites are cytotoxic to mammalian cells therefore making them important drug leads for cancer therapy.<sup>[68]</sup> The eukaryotic cytoskeleton is the main target by interference with the polymerization of actin (chondramides, chivosazols and rhizopodin) <sup>[79-81]</sup> and tubulin (disorazols, tubulysins, and epothilone) <sup>[82,83]</sup> as the most observed modes of action. At the forefront of myxobacterial chemotherapeutic agents is epothilone which is active against multidrug resistant cancer cell lines. Epothilone has a similar mode of action as paclitaxel, one of the most important anticancer to date and can therefore be used for the treatment of paclitaxel-resistant tumors.<sup>[84,85]</sup> Unlike the disorazols and tubulysins which promote tubulin disassembly, epothilone's mode of action is the stabilization of tubulin polymerization. The semi-synthetic analogue ixabepilone (Ixempra®) has already been approved by the US Food and Drug Administration (FDA) for the treatment of breast cancer.<sup>[86]</sup>

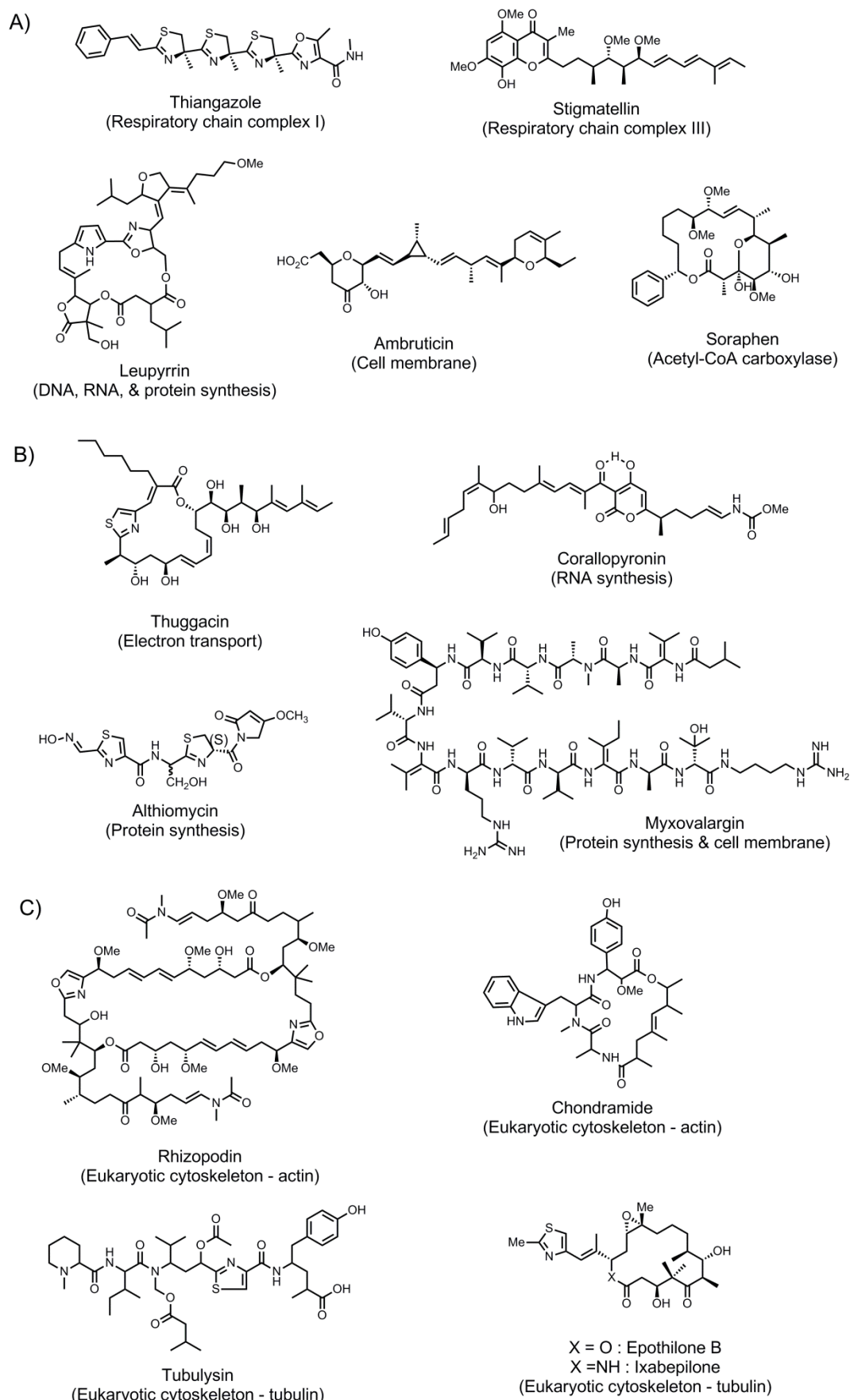
Similar to what is observed from actinomycetes, availability of genome sequence information revealed that the biosynthetic potential of myxobacteria to produce secondary metabolites is much higher than previously anticipated, with the number of putative products significantly higher than that what has been detected so far from sequenced strains.<sup>[35]</sup> To date, there are 6 published myxobacterial genomes (Table 1.1). The genomes are GC-rich and, with the exception of the genome size of *Anaeromyxobacter dehalogenans*, the largest among bacteria.<sup>[67]</sup>

## 1 Introduction

**Table 1.1:** Selected features of myxobacterial genome

Strain	Genome Size (Mbp)	GC content	Total number of ORFs	Number of ORFs with assigned functions	Ref.
<i>Sorangium cellulosum</i> So ce56	13.0	71.4%	9367	4895	[87]
<i>Stigmatella aurantiaca</i> DW4/3-1	10.2	67.5%	8586	4271	[88]
<i>Myxococcus xanthus</i> DK1622	9.1	68.9%	7388	3608	[89]
<i>Anaeromyxobacter dehalogenans</i> 2CP-C	5.0	74.9%	4287	3051	[90]
<i>Myxococcus fulvus</i> HW-1	9.0	70.6%	7285	2630	[91]
<i>Haliangium ochraceum</i> SMP-2 <sup>T</sup>	9.5	69.5%	6898	4284	[92]

Annotation of the genomes revealed that the strains *Sorangium cellulosum* So ce56, *Stigmatella aurantiaca* DW4/3-1, and *Myxococcus xanthus* DK1622 harbor multiple biosynthetic gene clusters encoding PKS, NRPS, and hybrid PKS/NRPS systems for which only a handful have been correlated to their secondary metabolite counterparts (Fig. 1.3).<sup>[35]</sup> Myxobacterial natural product discovery by genome mining is not straightforward. In terms of secondary metabolism, myxobacteria are biosynthetic nonconformists, breaking the rules of textbook assembly logic.<sup>[93]</sup> For example, in the biosynthesis of myxochromide S, module skipping and iterative PKS modules were identified.<sup>[43]</sup> The DH domain of AufD and KR domain of AufC of the aurafuron biosynthetic pathway are left unused during biosynthesis although the catalytic functions are theoretically intact.<sup>[94]</sup> The recent completion of the aurachin biosynthetic pathway showed that the essential catalytic functions are not encoded by genes in just one locus but are actually spread over three distinct loci in the genome.<sup>[95]</sup>



**Figure 1.8:** Bioactive myxobacterial natural products with A) antifungal, B) antibacterial, and C) cytotoxic activity. Biological targets are in parenthesis.

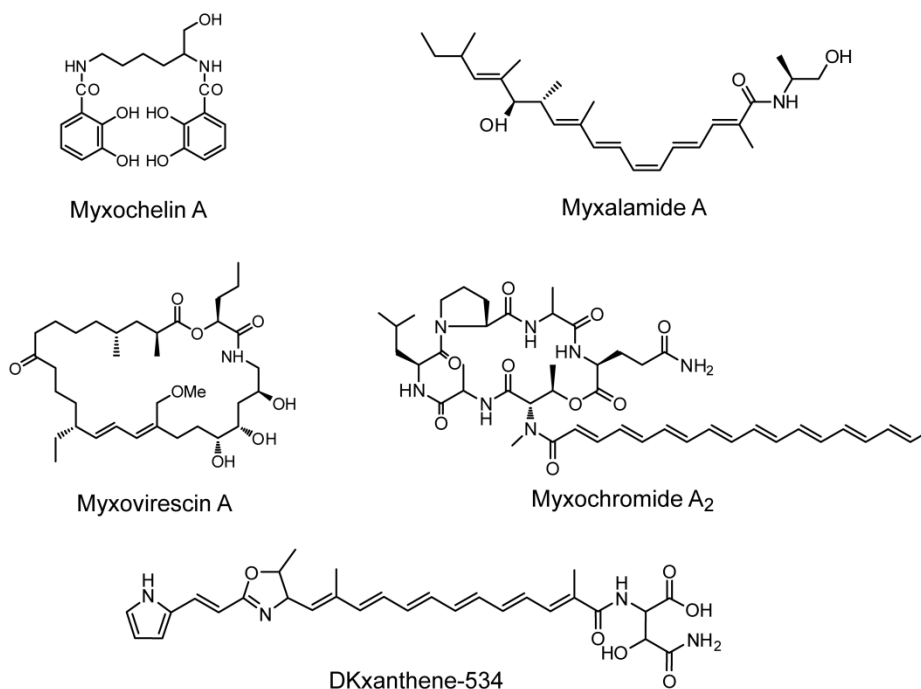
To further complicate matters, development of tools to genetically manipulate myxobacteria is a tedious process. Myxobacteria harbor an inherent resistance to commonly used antibiotics which limits the selection markers that can be used. Certain isolates initially grow as clumps or flakes and require a series of re-culturing before they grow homogeneously. Generally, genetic tools have to be tailor-made for each strain as protocols established for one strain often are not adaptable to another one, even in cases where the strains are closely phylogenetically related.<sup>[68;96]</sup>

In spite of what might seem to be a arduous task, by 2009, 25 complete biosynthetic gene clusters have been identified and published from myxobacteria. Since then 8 more have been discovered and characterized, two of which will be presented in this work.

### 1.5 The biosynthetic potential of *Myxococcus xanthus*

*Myxococcus xanthus* is a myxobacterial species of the suborder Cystobacterineae that subsist on the dung pellets of herbivores.<sup>[61]</sup> *M. xanthus* DK1622 is the central and best studied object of myxobacterial research as it is an excellent model system for investigating gliding motility, intercellular interaction, and multicellular organizations in microbial communities.<sup>[97;98]</sup> DK1622 has one of the largest microbial genomes with a size of about 9.14 Mbp.<sup>[89]</sup> This is partially ascribed to the large number of genes required to carry out its complex life cycle and signaling mechanisms.<sup>[99]</sup>

Prior to genome sequencing, *M. xanthus* DK1622 was not known to produce any secondary metabolites. Upon closer examination of the genome sequence, it turned out that 8.5% of its genome is dedicated to biosynthetic gene clusters (18 in total) encoding PKS, NRPS, and NRPS/PKS hybrid systems. Since then, five NRPS/PKS-derived natural product families from DK1622 have been identified: the myxalamids (*mx*<sub>A</sub>), myxochelins (*mch*<sub>A</sub>), myxochromides (*mx*<sub>C</sub>), myxovirescins (*ta*), and DKxanthenes (*dkx*). (Figure 1.9)



**Figure 1.9:** Representative structures of the *M. xanthus* DK1622 compound families

The *mx<sub>a</sub>*, *mx<sub>c</sub>*, and *mch<sub>A</sub>* gene clusters were easily dereplicated as their biosynthetic pathways have been previously identified and characterized from *S. aurantiaca* strains.<sup>[88]</sup> The myxovirescin *ta* gene cluster was identified by BLAST sequence analysis of a PKS/NRPS sequence fragment obtained from *M. xanthus* ER-1.<sup>[100]</sup> Due to the yellow/tan phase variance exhibited by the strain DK1622, the gene cluster associated with the yellow pigment DKxanthene was initially identified by transposon mutagenesis in the stably yellow strain DK1050. The sequence fragments obtained after vector recovery were used to identify the *dkx* locus in DK1622.<sup>[101]</sup>

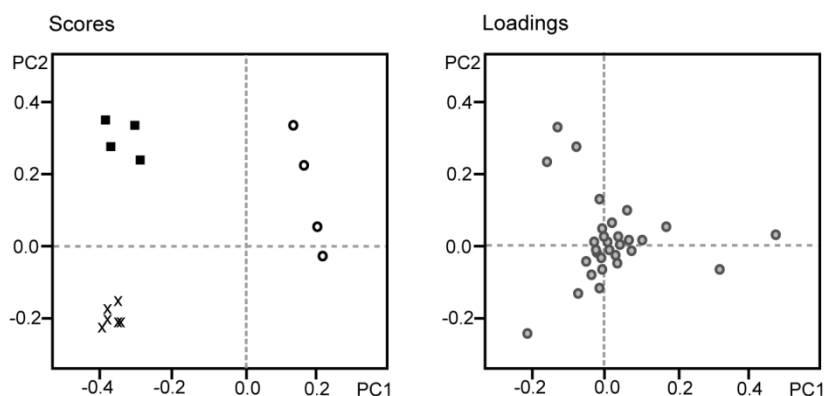
To determine the associated products of the remaining 13 unassigned biosynthetic pathways, each of the gene clusters was targeted by TM-MP. The manual comparison of wildtype and mutant profiles of the LC-MS data yielded no result.

The question remained: is it possible to detect new compounds from *M. xanthus*? A metabolomics-guided analysis of extracts from 98 *M. xanthus* isolates suggested the inherent potential. By using advanced analytics (fast separation UPLC, electrospray ionization (ESI) and high resolution time-of-flight mass spectrometry) coupled to multivariate statistics by principal component analysis, 37 nonubiquitous compounds were identified.<sup>[102]</sup>

## 1.6 Principal Component Analysis

Deciphering the differences among complex LC-MS data sets is not always easily accomplished particularly if the distinction of interest is subtle and is buried underneath the rich information. The task then requires statistical tools like principal component analysis to dig deep into the data to expose the differences.

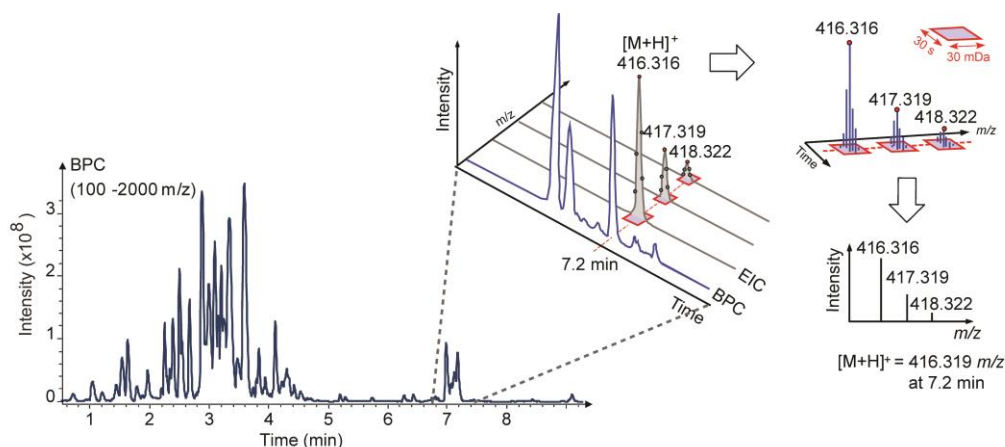
Principal component analysis (PCA) is an established multivariate data mining tool which summarizes observations represented by inter-correlated quantitative dependent variables. Important information is highlighted and compressed into a new set of orthogonal variables called principal components. The similarities, trends, and outliers are then pictorially represented by two dimensional maps called scores and loadings plots.<sup>[103]</sup>



**Figure 1.10:** Scores and loadings plot. The scores plot reflects the clustering of similar samples whereas the loadings plot provides the clustering of observations which underpin the grouping pattern of the samples.

PCA is a widely used unsupervised algorithm in systems biology and metabolomics to analyze complex NMR, GC-MS, CE-MS, and HPLC-MS data sets.<sup>[104]</sup> When analyzing metabolomic HPLC-MS data with PCA, the data set has to be reduced to discrete metabolites represented by mass-to-charge and retention time ( $m/z$ -RT) pairs and described quantitatively by the intensity.<sup>[105]</sup> The dimensionality can be further reduced if related signals (i.e. related by isotope pattern and charge states) are compiled into one discrete entity referred to as molecular feature (formerly termed compound bucket) represented by one  $m/z$ -RT pair and described by a representative peak intensity, for example the intensity of the monoisotopic peak (Figure 1.11).<sup>[106]</sup>





**Figure 1.11:** The finding molecular feature (FMF) algorithm identifies isotope peaks and charge states from a common parent ion in subsequent spectra and combines this information into one compound, thereby removing redundant information from a dataset.<sup>[106]</sup>

Determining the phenotype of the targeted mutagenesis on the 13 unassigned biosynthetic pathways of *M. xanthus* is a well suited problem that can be addressed by PCA as the difference between wildtype and mutant strains could ultimately be a single, low intensity peak in an information-rich LC-MS data set.

## 1.7 Outline of the study

The main topic of this work is to associate biosynthetic pathways identified in *Myxococcus xanthus* to their corresponding secondary metabolite counterparts. These biosynthetic pathways are PKSs, NRPSs, and hybrid synthetases encoded in distinct genetic loci in the *M. xanthus* genome.

The first part delves into the identification of the associated products of the 13 unassigned pathways in *M. xanthus* DK1622. Evaluation at the transcriptomic and proteomic levels asserts that the remaining 13 unassigned biosynthetic gene clusters are active under standard conditions for the cultivation of DK1622.<sup>[107;108]</sup> It is presumed that the low abundance of the corresponding compounds in general may complicate their detection, thereby impeding the identification of new structures. By coupling gene inactivation studies with extensive ultra-high-performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS) and downstream comprehensive comparison of secondary metabolomes by PCA and hypothesis driven statistics, three novel *M. xanthus* compound classes were simultaneously correlated to their biosynthetic pathways. In the context of pathways to products, the comprehensive secondary metabolome approach is an extension of the TM-MP method with the refinement that

statistics is capable of finding the subtle differences between wildtype and mutant metabolite profiles. After optimization of production, myxoprincomide-c506 was fully characterized by NMR and tandem MS experiments. Myxoprincomide, a linear peptide bearing unusual structural features, is produced by the NRPS/PKS hybrid system encoded by *mxp* (MXAN\_3779) biosynthetic machinery.

The second part expounds on the Group I myxoprincomide compound class. Large scale cultivation for the isolation of myxoprincomide-c506 unearthed multiple derivatives ranging from shorter products to larger myxoprincomide variants. Concurrently, a heterofunctional homolog of the *mxp* gene was identified in a related strain *M. xanthus* DK897, yielding a different set of myxoprincomide structures (Group II). Supported by feeding studies, spectroscopic and spectrometric data, the structures of the Group I myxoprincomide family and a biosynthetic hypothesis are presented.

The third section deals with the identification and characterization of the althiomycin biosynthetic gene cluster in *M. xanthus* DK897. Althiomycin, (Fig. 1.8 B) a heavily modified pentapeptide (MF: C<sub>18</sub>H<sub>17</sub>N<sub>5</sub>O<sub>6</sub>S<sub>2</sub>, MW : 439 g/mol), was first isolated in 1957 from *Streptomyces althioticus*.<sup>[109]</sup> It is a broad spectrum antibiotic that inhibits prokaryotic protein synthesis at the peptidyltransfer stage but exhibits no such effects in mammalian cells.<sup>[77]</sup> SAR studies revealed that the configuration of the oxime group, the configuration at C4 of the thiazoline ring, the methoxypyrrolinone and, to a lesser extent, the hydroxymethyl group are relevant for activity.<sup>[110;111]</sup> A previous screening of myxobacteria for novel antibiotics revealed that members of the genera *Cystobacter* and *Myxococcus* produce althiomycin.<sup>[112]</sup> Although there are other examples described (e.g. saframycin), it is rather unusual to find a myxobacterial secondary metabolite to be also produced by a streptomycete.<sup>[113]</sup> Through genome mining, targeted gene inactivation, and heterologous expression, the biosynthetic pathway was identified and a model for the assembly of althiomycin was proposed.

## 1.8 References

- [1.] D. J. Newman, G. M. Cragg, K. M. Snader, *Nat.Prod.Rep.* **2000**, 17 215-234.
- [2.] F. E. Koehn, G. T. Carter, *Nat.Rev.Drug Discovery* **2005**, 4 206-220.
- [3.] D. J. Newman, *J.Med.Chem.* **2008**, 51 2589-2599.
- [4.] D. J. Newman, G. M. Cragg, *J.Nat.Prod.* **2007**, 70 461-477.
- [5.] D. J. Newman, G. M. Cragg, *Curr.Opin.Investig.Drugs* **2009**, 10 1280-1296.
- [6.] S. Donadio, S. Maffioli, P. Monciardini, M. Sosio, D. Jabes, *J.Antibiot.* **2010**, 63 423-430.
- [7.] G. T. Carter, *Nat.Prod.Rep.* **2011**, 28 1783-1789.
- [8.] C. T. Walsh, M. A. Fischbach, *J.Am.Chem.Soc.* **2010**, 132 2469-2493.
- [9.] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem Rev* **1997**, 97 2651-2674.
- [10.] D. Schwarzer, R. Finking, M. A. Marahiel, *Nat.Prod.Rep.* **2003**, 20 275-287.
- [11.] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem.Biol.* **1999**, 6 493-505.
- [12.] J. J. May, N. Kessler, M. A. Marahiel, M. T. Stubbs, *P.Natl.Acad.Sci.USA* **2002**, 99 12120-12125.
- [13.] B. O. Bachmann, J. Ravel, *Methods in Enzymology* **2009**, 458 181-217.
- [14.] M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch, O. Kohlbacher, *Nucleic Acids Res.* **2011**, 39 W362-W367.
- [15.] C. Rausch, T. Weber, O. Kohlbacher, W. Wohlleben, D. H. Huson, *Nucleic Acids Res.* **2005**, 33 5799-5808.
- [16.] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem.Biol.* **1996**, 3 923-936.
- [17.] L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry-US* **1998**, 37 1585-1595.
- [18.] S. A. Sieber, M. A. Marahiel, *Chem Rev* **2005**, 105 715-738.
- [19.] K. L. Dunbar, J. O. Melby, D. A. Mitchell, *Nat Chem Biol Nat Chem Biol* **2012**, 8 569-575.
- [20.] L. Du, L. Lou, *Nat.Prod.Rep.* **2010**, 27 255-278.
- [21.] B. S. Moore, C. Hertweck, *Nat.Prod.Rep.* **2002**, 19 70-99.
- [22.] M. A. Fischbach, C. T. Walsh, *Chem.Rev.* **2006**, 106 3468-3496.

- [23.] G. Yadav, R. S. Gokhale, D. Mohanty, *Journal of Molecular Biology* **2003**, 328 335-363.
- [24.] G. Yadav, R. S. Gokhale, D. Mohanty, *Nucleic Acids Res.* **2003**, 31 3654-3658.
- [25.] F. Del Vecchio, H. Petkovic, S. G. Kendrew, L. Low, B. Wilkinson, R. Lill, J. Cortes, B. A. Rudd, J. Staunton, P. F. Leadlay, *J.Ind.Microbiol.Biotechnol.* **2003**, 30 489-494.
- [26.] H. Petkovic, A. Sandmann, L. R. Challis, H. J. Hecht, B. Silakowski, L. Low, N. Beeston, E. Kuscer, J. Garcia-Bernardo, P. F. Leadlay, S. G. Kendrew, B. Wilkinson, R. Müller, *Org.Biomol.Chem.* **2008**, 6 500-506.
- [27.] K. J. Weissman, *Ernst Schering Res Found Workshop* **2005**, 43-78.
- [28.] C. Hertweck, A. Luzhetskyy, Y. Rebets, A. Bechthold, *Nat.Prod.Rep.* **2007**, 24 162-190.
- [29.] M. B. Austin, J. P. Noel, *Nat.Prod.Rep.* **2003**, 20 79-110.
- [30.] J. M. Winter, S. Behnken, C. Hertweck, *Curr.Opin.Chem.Biol.* **2010**.
- [31.] G. Lang, N. A. Mayhudin, M. I. Mitova, L. Sun, S. S. van der, J. W. Blunt, A. L. Cole, G. Ellis, H. Laatsch, M. H. Munro, *J.Nat.Prod.* **2008**, 71 1595-1599.
- [32.] K. F. Nielsen, M. Mansson, C. Rank, J. C. Frisvad, T. O. Larsen, *J.Nat.Prod.* **2011**, 74 2338-2348.
- [33.] M. Nett, H. Ikeda, B. S. Moore, *Nat.Prod.Rep.* **2009**, 26 1362-1384.
- [34.] J. A. Kalaitzis, F. M. Lauro, B. A. Neilan, *Nat.Prod.Rep.* **2009**, 26 1447-1465.
- [35.] S. C. Wenzel, R. Müller, *Nat.Prod.Rep.* **2009**, 26 1385-1407.
- [36.] H. Gross, J. E. Loper, *Nat.Prod.Rep.* **2009**, 26 1408-1446.
- [37.] S. Donadio, P. Monciardini, M. Sosio, *Nat.Prod.Rep.* **2007**, 24 1073-1109.
- [38.] H. Gross, V. O. Stockwell, M. D. Henkels, B. Nowak-Thompson, J. E. Loper, W. H. Gerwick, *Chem.Biol.* **2007**, 14 53-63.
- [39.] A. Pospiech, B. Cluzel, H. Bietenhader, T. Schupp, *Microbiology* **1995**, 141 1793-1803.
- [40.] B. Silakowski, H. U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blöcker, G. Höfle, S. Beyer, R. Müller, *J.Biol.Chem.* **1999**, 274 37391-37399.
- [41.] L. Tang, S. Shah, L. Chung, J. Carney, L. Katz, C. Khosla, B. Julien, *Science* **2000**, 287 640-642.
- [42.] A. Sandmann, F. Sasse, R. Müller, *Chem.Biol.* **2004**, 11 1071-1079.

- [43.] S. C. Wenzel, B. Kunze, G. Höfle, B. Silakowski, M. Scharfe, H. Blöcker, R. Müller, *ChemBioChem* **2005**, 6 375-385.
- [44.] N. Ziemert, K. Ishida, A. Liaimer, C. Hertweck, E. Dittmann, *Angew.Chem.Int.Ed.* **2008**, 47 7756-7759.
- [45.] W. K. W. Chou, I. Fanizza, T. Uchiyama, M. Komatsu, H. Ikeda, D. E. Cane, *J.Am.Chem.Soc.* **2010**, 132 8850-8851.
- [46.] S. C. Wenzel, F. Gross, Y. Zhang, J. Fu, F. A. Stewart, R. Müller, *Chem.Biol.* **2005**, 12 349-356.
- [47.] O. Perlova, J. Fu, S. Kuhlmann, D. Krug, F. Stewart, Y. Zhang, R. Müller, *Appl.Environ.Microbiol.* **2006**, 72 7485-7494.
- [48.] J. Fu, X. Bian, S. Hu, H. Wang, F. Huang, P. M. Seibert, A. Plaza, L. Xia, R. Müller, A. F. Stewart, Y. Zhang, *Nat.Biotechnol.* **2012**, 30 440-446.
- [49.] H. B. Bode, B. Bethe, R. Höfs, A. Zeeck, *ChemBioChem* **2002**, 3 619-627.
- [50.] K. Scherlach, C. Hertweck, *Org.Biomol.Chem* **2009**, 7 1753-1760.
- [51.] K. Scherlach, C. Hertweck, *Org.Biomol.Chem.* **2006**, 4 3517-3520.
- [52.] S. Bergmann, A. N. Funk, K. Scherlach, V. Schroeckh, E. Shelest, U. Horn, C. Hertweck, A. A. Brakhage, *Appl.Environ.Microbiol.* **2010**, 76 8143-8149.
- [53.] Y. M. Chiang, E. Szewczyk, A. D. Davidson, N. Keller, B. R. Oakley, C. C. C. Wang, *J.Am.Chem.Soc.* **2009**, 131 2965-2970.
- [54.] S. B. Bumpus, B. S. Evans, P. M. Thomas, I. Ntai, N. L. Kelleher, *Nat.Biotechnol.* **2009**.
- [55.] B. S. Evans, I. Ntai, Y. Q. Chen, S. J. Robinson, N. L. Kelleher, *J.Am.Chem.Soc.* **2011**, 133 7316-7319.
- [56.] Y. Chen, I. Ntai, K. S. Ju, M. Unger, L. Zamdborg, S. J. Robinson, J. R. Doroghazi, D. P. Labeda, W. W. Metcalf, N. L. Kelleher, *J Proteome Res J Proteome Res* **2012**, 11 85-94.
- [57.] J. Meier, S. Niessen, H. Hoover, T. L. Foley, B. F. Cravatt, M. D. Burkart, *ACS Chem.Biol.* **2009**.
- [58.] W. P. Heal, T. H. T. Dang, E. W. Tate, *Chem.Soc.Rev.* **2011**, 40 246-257.
- [59.] X. Qu, C. Lei, Liu W., *Angew.Chem.Int.Ed.* **2011**, 50 9651-9654.
- [60.] R. O. Garcia, D. Krug, R. Müller, in *Methods in Enzymology* **2009**, p. pp. 59-91.
- [61.] W. Dawid, *FEMS Microbiol.Rev.* **2000**, 24 403-427.
- [62.] W. Dawid, C. A. Gallikowski, P. Hirsch, *Polarforschung* **1988**, 58 271-278.

- [63.] R. Fudou, Y. Jojima, T. Iizuka, S. Yamanaka, *Journal of General and Applied Microbiology* **2002**, 48 109-115.
- [64.] K. Gerth, R. Müller, *Environ.Microbiol.* **2005**, 7 874-880.
- [65.] L. Shimkets, M. Dworkin, H. Reichenbach, in *The Prokaryotes* ,Vol. 7, (Ed.: M. Dworkin), Springer, Berlin **2006**, p. pp. 31-115.
- [66.] K. J. Weissman, R. Müller, *Nat.Prod.Rep.* **2010**, 27 1276-1295.
- [67.] S. C. Wenzel, R. Müller, *Curr Opin Drug Discov Devel* **2009**, 12 220-230.
- [68.] K. J. Weissman, R. Müller, *Bioorganic & Medicinal Chemistry* **2009**, 17 2121-2136.
- [69.] K. Gerth, S. Pradella, O. Perlova, S. Beyer, R. Müller, *J.Biotechnol.* **2003**, 106 233-253.
- [70.] Y. Xiao, X. M. Wei, R. Ebright, D. Wall, *J.Bacteriol.* **2011**, 193 4626-4633.
- [71.] Esposti M.D., *Biochim.Biophys.Acta* **1998**, 1364 222-235.
- [72.] H. B. Bode, H. Irschik, S. C. Wenzel, H. Reichenbach, R. Müller, G. Höfle, *J.Nat.Prod.* **2003**, 66 1203-1206.
- [73.] F. Sasse, H. Steinmetz, G. Höfle, H. Reichenbach, *J.Antibiot.* **1995**, 48 21-25.
- [74.] P. Knauth, H. Reichenbach, *J.Antibiot.* **2000**, 53 1182-1190.
- [75.] H. Steinmetz, H. Irschik, B. Kunze, H. Reichenbach, G. Höfle, R. Jansen, *Chemistry* **2007**, 13 5822-5832.
- [76.] J. Mukhopadhyay, K. Das, S. Ismail, D. Koppstein, M. Jang, B. Hudson, S. Sarafianos, S. Tuske, J. Patel, R. Jansen, H. Irschik, E. Arnold, R. H. Ebright, *Cell* **2008**, 135 295-307.
- [77.] H. Fujimoto, T. Kinoshita, H. Suzuki, H. Umezawa, *J.Antibiot.* **1970**, 23 271-275.
- [78.] H. Irschik, H. Reichenbach, *J.Antibiot.* **1985**, 38 1237-1245.
- [79.] F. Sasse, B. Kunze, T. M. Gronewold, H. Reichenbach, *J.Natl.Cancer Inst.* **1998**, 90 1559-1563.
- [80.] T. M. Gronewold, F. Sasse, H. Lunsdorf, H. Reichenbach, *Cell Tissue Res* **1999**, 295 121-129.
- [81.] R. Diestel, H. Irschik, R. Jansen, M. W. Khalil, H. Reichenbach, F. Sasse, *Chembiochem.* **2009**.
- [82.] Y. A. Elnakady, F. Sasse, H. Lünsdorf, H. Reichenbach, *Biochem.Pharmacol.* **2004**, 67 927-935.
- [83.] M. W. Khalil, F. Sasse, H. Lünsdorf, Y. A. Elnakady, H. Reichenbach, *ChemBioChem* **2006**, 7 678-683.

- 
- [84.] J. Mulzer, *The Epothilones - An Outstanding Family of Anti-Tumour Agents: From Soil to the Clinic*, (Ed.: J. Mulzer) Springer, New York **2009**.
  - [85.] H. Reichenbach, G. Höfle, *Drugs in R&D* **2008**, 9 1-10.
  - [86.] P. Fumoleau, B. Coudert, N. Isambert, E. Ferrant, *Ann.Oncol.* **2007**, 18 v9-v15.
  - [87.] S. Schneiker, O. Perlova, O. Kaiser, K. Gerth, A. Alici, M. O. Altmeyer, D. Bartels, T. Bekel, S. Beyer, E. Bode, H. B. Bode, C. J. Bolten, J. V. Choudhuri, S. Doss, Y. A. Elnakady, B. Frank, L. Gaigalat, A. Goesmann, C. Groeger, F. Gross, L. Jelsbak, L. Jelsbak, J. Kalinowski, C. Kegler, T. Knauber, S. Konietzny, M. Kopp, L. Krause, D. Krug, B. Linke, T. Mahmud, R. Martinez-Arias, A. C. McHardy, M. Merai, F. Meyer, S. Mormann, J. Munoz-Dorado, J. Perez, S. Pradella, S. Rachid, G. Raddatz, F. Rosenau, C. Rückert, F. Sasse, M. Scharfe, S. C. Schuster, G. Suen, A. Treuner-Lange, G. J. Velicer, F. J. Vorhölter, K. J. Weissman, R. D. Welch, S. C. Wenzel, D. E. Whitworth, S. Wilhelm, C. Wittmann, H. Blöcker, A. Pühler, R. Müller, *Nat.Biotechnol.* **2007**, 25 1281-1289.
  - [88.] C. M. Ronning, W. C. Nierman, in *Myxobacteria: Multicellularity and differentiation* (Ed.: D. Whitworth), ASM Press, Chicago **2007**, p. pp. 285-298.
  - [89.] B. S. Goldman, W. C. Nierman, D. Kaiser, S. C. Slater, A. S. Durkin, J. Eisen, C. M. Ronning, W. B. Barbazuk, M. Blanchard, C. Field, C. Halling, G. Hinkle, O. Iartchuk, H. S. Kim, C. Mackenzie, R. Madupu, N. Miller, A. Shvartsbeyn, S. A. Sullivan, M. Vaudin, R. Wiegand, H. B. Kaplan, *P.Natl.Acad.Sci.USA* **2006**, 103 15200-15205.
  - [90.] S. H. Thomas, R. D. Wagner, A. K. Arakaki, J. Skolnick, J. R. Kirby, L. J. Shimkets, R. A. Sanford, F. E. Löffler, *PLoS ONE*. **2008**, 3 e2103.
  - [91.] Z. F. Li, X. Li, H. Liu, X. Liu, K. Han, Z. H. Wu, W. Hu, F. F. Li, Y. Z. Li, *J.Bacteriol.* **2011**, 193 5015-5016.
  - [92.] N. Ivanova, C. Daum, E. Lang, B. Abt, M. Kopitz, E. Saunders, A. Lapidus, S. Lucas, T. G. Del Rio, M. Nolan, H. Tice, A. Copeland, J. F. Cheng, F. Chen, D. Bruce, L. Goodwin, S. Pitluck, K. Mavromatis, A. Pati, N. Mikhailova, A. Chen, K. Palaniappan, M. Land, L. Hauser, Y. J. Chang, C. D. Jeffries, J. C. Detter, T. Brettin, M. Rohde, M. Goker, J. Bristow, V. Markowitz, J. A. Eisen, P. Hugenholtz, N. C. Kyrpides, H. P. Klenk, *Standards in Genomic Sciences* **2010**, 2 96-106.
  - [93.] S. C. Wenzel, R. Müller, *Nat.Prod.Rep.* **2007**, 24 1211-1224.
  - [94.] B. Frank, S. C. Wenzel, H. B. Bode, M. Scharfe, H. Blöcker, R. Müller, *Journal of Molecular Biology* **2007**, 374 24-38.
  - [95.] D. Pistorius, Y. Li, A. Sandmann, K. J. Weissman, R. Müller, *Molecular Biosystems* **2011**, 7 3308-3315.
  - [96.] M. Kopp, H. Irschik, F. Gross, O. Perlova, A. Sandmann, K. Gerth, R. Müller, *J.Biotechnol.* **2004**, 107 29-40.

- [97.] B. Nan, D. R. Zusman, *Annu.Rev.Genet.* **2011**, 45 21-39.
- [98.] G. J. Velicer, M. Vos, *Annual Review of Microbiology* **2009**, 63 599-623.
- [99.] D. Kaiser, *Annual Review of Microbiology* **2004**, 58 75-98.
- [100.] V. Simunovic, J. Zapp, S. Rachid, D. Krug, P. Meiser, R. Müller, *ChemBioChem* **2006**, 7 1206-1220.
- [101.] P. Meiser, K. J. Weissman, H. B. Bode, D. Krug, J. S. Dickschat, A. Sandmann, R. Müller, *Chem.Biol.* **2008**, 15 771-781.
- [102.] D. Krug, G. Zurek, O. Revermann, M. Vos, G. J. Velicer, R. Müller, *Appl.Environ.Microbiol.* **2008**, 74 3058-3068.
- [103.] H. Abdi, L. J. Williams, *WIREs Computational Statistics* **2010**, 2 433-459.
- [104.] V. Shulaev, *Briefings in Bioinformatics* **2006**, 7 128-139.
- [105.] A. Nordström, in *Metabolomics, Metabonomics and Metabolite Profiling* Ed.: W. J. Griffiths) **2008**, p. pp. 273-294.
- [106.] D. Krug, G. Zurek, B. Schneider, R. Garcia, R. Müller, *Anal.Chim.Acta* **2008**, 624 97-106.
- [107.] H. B. Bode, M. W. Ring, G. Schwär, M. O. Altmeyer, C. Kegler, I. R. Jose, M. Singer, R. Müller, *ChemBioChem* **2009**, 10 128-140.
- [108.] C. Schley, M. O. Altmeyer, R. Swart, R. Müller, C. G. Huber, *J Proteome Res J Proteome Res* **2006**, 5 2760-2768.
- [109.] H. Yamaguchi, Y. Nakayama, K. Takeda, K. Tawara, K. Maeda, T. Takeuchi, H. Umezawa, *J.Antibiot.* **1957**, 10 195-200.
- [110.] K. Inami, T. Shiba, *Bull.Chem.Soc.Jpn.* **1986**, 59 2185-2189.
- [111.] P. Zarantonello, C. P. Leslie, R. Ferritto, W. M. Kazmierski, *Bioorg Med Chem Lett Bioorg Med Chem Lett* **2002**, 12 561-565.
- [112.] B. Kunze, H. Reichenbach, H. Augustiniak, G. Höfle, *J.Antibiot.* **1982**, 35 635-636.
- [113.] H. Reichenbach, G. Höfle, in *Drug Discovery from Nature* Eds.: S. Grabley, R. Thiericke), Springer, Berlin **1999**, p. pp. 149-179.



## 1.9 Appendix I – List of myxobacterial compound families

Ajudazol	Cittilin	Icumazole	Nannochelin	Spirangien
Althiomycin	Corallopyronins	Indolylacetaldoxim	Noricumazole	Stigmatellin
Ambruticin	Crocacin	Jerangolid	Pellasoren	Sulasoren
Angiolam	Crocapeptin	Kulkenon	PentacaronsÄure	Tartrolon
Antalam	Cruentaren	Leupyrrin	Phenalamide	Terrestribisamid
Antalid	Cyrmenin	Maltepolide	Phenoxan	Thiangazole
Apicularen	Dawenol	Maracen	Pyrrolnitrin	Thuggacin
Archazolid	Disorazole	Maracin	Pyrronazol	Tichunal
Argyrim	DKxanthene	Marinoquinolines	Ratjadon	Tubulysin
Aurachin	Elansolide	Melithiazol	Rhizopodin	Tuscolid
Aurafuron	Eliamid	Myxalamid	Ripostatin	Tuscoron
Carolacton	Epothilone	Myxochelin	Roimatacen	Varnalamide
Cerebrosid	Etnangien	Myxochromide	Saframycin	WAP
Chivosazole	Gephyronic acid	Myxoprincomide	Socein	
Chlorotonil	Germacran	Myxopyronins	Soracumen	
Chlorsoraphene	Haprolid	Myxothiazole	Sorangicin	
Chondramide	Holmenon	Myxoalargins	Sorangiolid	
Chondrochloren	Hymeron	Myxovirescin	Soraphen	

## 2 Materials and Methods

### 2.1 Chemicals

All chemicals used in this work were obtained in analytical grade from the corresponding manufacturer.

**Table 2.1:** Chemicals used and their sources

Chemical	Manufacturer
Acetonitrile	Sigma Aldrich
Trifluoroacetic acid	
Ethanol	
Tris base	
XAD-16 resin	
Bromophenol blue	
Nicotinamide	
<i>p</i> -Aminobenzoic acid	
MnCl <sub>2</sub> • 4H <sub>2</sub> O	
DMSO	
Ethyl Acetate	Fischer Scientific
Methanol	
2-Propanol	
Formic Acid	
KOH	Fluka
HCl	
Amino acid standards	
Xylene cyanol	
FE(III)-EDTA	
EDTA	
Starch	
HEPES	
Kanamycin sulfate	
Roti®-Phenol/Chloroform/Isoamyl alcohol	
Potassium acetate	Carl Roth
Vitamin B12	
Ampicillin	
n-Butanol	
NaCl	VWR

<b>KH<sub>2</sub>PO<sub>4</sub></b>	
<b>n-Heptane</b>	ZChL
<b>Glacial acetic acid</b>	
<b>Glycerin</b>	
<b>ZnCl<sub>2</sub></b>	
<b>Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub> • 10H<sub>2</sub>O</b>	
<b>(NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub> • 4H<sub>2</sub>O</b>	
<b>K<sub>2</sub>HPO<sub>4</sub></b>	Grüssing
<b>MOPS</b>	MP Biomedicals
<b>SDS</b>	
<b>Biotin</b>	Merck
<b>CaCl<sub>2</sub> • 2H<sub>2</sub>O</b>	
<b>MgSO<sub>4</sub> • 7H<sub>2</sub>O</b>	
<b>NaOH</b>	
<b>Pyridoxal-HCl</b>	
<b>FeCl<sub>3</sub> • 6H<sub>2</sub>O</b>	
<b>Zeocin™ (100 mg/ml)</b>	Invitrogen
<b>Sephadex LH-20</b>	Amersham Biosciences
<b>5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal)</b>	Fermentas
<b>dATPs</b>	
<b>dCTPs</b>	
<b>dGTPs</b>	
<b>dTTPs</b>	
<b>Bacto Agar</b>	Becton Dickinson and Co., USA
<b>Casitone</b>	
<b>Tryptone</b>	
<b>Yeast Extract</b>	
<b>Baker's Yeast</b>	
<b>Casein Peptone Type S</b>	Marcor
<b>Agarose</b>	Biozym
<b>5-fluoro-DL-tryptophan</b>	Argos Organics
<b>Isotopically-labeled substrates</b>	
<b>L-alanine-2-d</b>	Isotec
<b>L-leucine-5,5,5-d<sub>3</sub></b>	Deutero
<b>L-leucine-2,3,3,4,5,5,5',5',5'-d<sub>10</sub></b>	Campro Scientific
<b>L-glutamic-2,3,3,4,4-d<sub>5</sub> acid</b>	Isotec
<b>L-serine- 2,3,3-d<sub>3</sub></b>	

## 2 Materials and Methods

L-serine- <sup>13</sup> C <sub>3</sub> , <sup>15</sup> N	CIL
L-cysteine- <sup>13</sup> C <sub>3</sub> , <sup>15</sup> N	Isotec
L-valine-d <sub>8</sub>	Isotec
L-threonine - <sup>13</sup> C <sub>4</sub> , <sup>15</sup> N	Deutero
L-phenylalanine- <sup>13</sup> C <sub>6</sub>	Isotec
L-histidine- <sup>13</sup> C <sub>6</sub>	euriso-top
L-methionine-(methyl-d <sub>3</sub> ),	CIL
L-methionine-(methyl- <sup>13</sup> C)	Campro Scientific
sodium 1- <sup>13</sup> C-acetate	CIL
sodium 2- <sup>13</sup> C-acetate	CIL
1,2- <sup>13</sup> C <sub>2</sub> -acetate	CIL
	CIL
Nα-(5-Fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA)	TCI
Nα-(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide (D-FDLA)	

## 2.2 Enzymes, kits, and markers

All enzymes, kits, and markers used in this work are listed in Table 2.2 and were used according to the manufacturers' protocols.

**Table 2.2:** Enzymes, kits and markers and their manufacturers

Enzyme	Manufacturer
<i>Taq</i> polymerase	Fermentas
Restriction endonucleases	
T4 DNA ligase	
Shrimp Alkaline Phosphatase	
GeneRuler™ 1 kb Plus DNA ladder	
GeneRuler™ 100bp Plus DNA ladder	
GeneJET™ Plasmid Miniprep Kit	
Phusion DNA polymerase	Finnzymes
TOPO TA Cloning ® Kit	Invitrogen
NucleoSpin® Gel and PCR Clean-up	Macherey & Nagel
Puregene Core Kit A	Qiagen
RNAse A	

## 2.3 Buffers and stock solutions

**Table 2.3:** Buffers and solutions for molecular biology

Buffers	Components	Amount
<b>Resuspension buffer (P1)</b>	Tris-HCl pH 8.0	0.61 g
	EDTA	0.29 g
	RNAse A	10 mg
	H <sub>2</sub> O	to 100 mL
<b>Lysis buffer (P2)</b>	NaOH	0.80 g
	SDS	1 g
	H <sub>2</sub> O	to 100 mL
<b>Neutralization buffer (P3)</b>	3 M CH <sub>3</sub> COOK	29.4 g
	H <sub>2</sub> O	to 100 mL
	glacial CH <sub>3</sub> COOH	to pH 5.5
<b>DNA loading dye</b>	Glycerin (87%)	3 mL
	Bromphenol blue	25 mg
	Xylene cyanol	25 mg
	H <sub>2</sub> O	to 10 mL
<b>TAE Buffer (1x)</b>	Tris base	2.42 g
	Acetic acid	572 µL
	0.5 M EDTA	1 mL
	H <sub>2</sub> O	to 500 mL
<b>0.5 M EDTA pH 8.0</b>	EDTA	18.61 g
	NaOH	2g
	H <sub>2</sub> O	to 100 mL
<b>50 % (v/v) Glycerol</b>	Glycerol	50 mL
	H <sub>2</sub> O	50 mL

## 2 Materials and Methods

**Table 2.4:** Buffers and solutions for media preparation. The following buffers were sterilized by autoclaving prior to use.

Buffers	Components	Amount
1 M Tris-HCl pH 8.0	Tris base	12.11 g
	HCl	to pH 8.0
	H <sub>2</sub> O	to 100 mL
0.8 M MgSO <sub>4</sub>	MgSO <sub>4</sub> • 7H <sub>2</sub> O,	19.68 g
	H <sub>2</sub> O	to 100 mL
1 M K <sub>2</sub> HPO <sub>4</sub> pH 7.6	1 M K <sub>2</sub> HPO <sub>4</sub>	86.6 ml
	1 M KH <sub>2</sub> PO <sub>4</sub>	13.4 ml
1 M KH <sub>2</sub> PO <sub>4</sub>	KH <sub>2</sub> PO <sub>4</sub>	13.61 g
	H <sub>2</sub> O	to 100 mL
1 M K <sub>2</sub> HPO <sub>4</sub>	K <sub>2</sub> HPO <sub>4</sub>	17.42 g
	H <sub>2</sub> O	to 100 mL
Fe(III)-EDTA (80 mg mL <sup>-1</sup> )	Fe(III)-EDTA	8 g
	H <sub>2</sub> O	to 100 mL
1 M MOPS pH 7.6	MOPS	20.93
	NaOH	to pH 7.6
	H <sub>2</sub> O	to 100 mL
Vitamin solution	Biotin	4 mg
	Sodium panthothenate	68 mg
	Nicotinamide	68 mg
	Pyridoxal-HCl	68 mg
	p-Aminobenzoic acid	14 mg
	Vitamin B12	15 mg
	dd H <sub>2</sub> O	3 mL
	Methanol	7 mL
Trace element solution	ZnCl <sub>2</sub>	40 mg
	FeCl <sub>3</sub> • 6H <sub>2</sub> O	200 mg
	CuCl <sub>2</sub> • 2H <sub>2</sub> O	10 mg
	MnCl <sub>2</sub> • 4H <sub>2</sub> O	10 mg
	Na <sub>2</sub> B <sub>4</sub> O <sub>7</sub> • 10H <sub>2</sub> O	10 mg
	(NH <sub>4</sub> ) <sub>6</sub> Mo <sub>7</sub> O <sub>24</sub> • 4H <sub>2</sub> O	10 mg
	dd H <sub>2</sub> O	to 1 L

## 2.4 Media

The following media were used in this work. For standard agar plate preparation, 15 g of agar were added per liter of liquid medium. For soft agar preparation, one liter of liquid medium was supplemented with 7.5 g agar.

### 2.4.1 LB Medium for *E. coli* cultivation

**Table 2.5:** Components of LB medium

Component	Amount
Tryptone	10 g
Yeast Extract	5 g
NaCl	5 g
H <sub>2</sub> O	to 1 L

### 2.4.2 Media for *M. xanthus* cultivation

The following media were used for the cultivation of *M. xanthus* during the course of this study. For large scale cultivation, the medium was supplemented with Fe(III)-EDTA to a final concentration of 80  $\mu\text{g mL}^{-1}$  to reduce the production of myxochelin.

**Table 2.6:** Media used for the cultivation of *M. xanthus*

Medium	Components	Amount	Final pH
CTT	Casitone	10 g	7.6
	1 M Tris-HCl pH 8.0	10 mL	
	0.8 M MgSO <sub>4</sub>	10 mL	
	1M K <sub>2</sub> HPO <sub>4</sub> (pH 7.6)	1 mL	
	H <sub>2</sub> O	to 1 L	
TS	Tryptone	10 g	7.2
	MgSO <sub>4</sub> • 7H <sub>2</sub> O	2 g	
	Starch	4 g	
	HEPES	1.19 g	
	H <sub>2</sub> O	to 1 L	
PS	Peptone Marcor S	10 g	7.0
	MgSO <sub>4</sub> • 7H <sub>2</sub> O	2 g	
	Starch	4 g	
	HEPES	1.19 g	
	H <sub>2</sub> O	to 1 L	
VY/2	Baker's yeast	5 g	7.0

## 2 Materials and Methods

	HEPES	1.19	
	CaCl <sub>2</sub> • 2H <sub>2</sub> O	0.5	
	H <sub>2</sub> O	to 1 L	
CYE	Casitone	10 g	7.6
	Yeast extract	5 g	
	1 M MOPS, pH 7.6	10 mL	
	MgSO <sub>4</sub> • 7H <sub>2</sub> O	1 g	
	H <sub>2</sub> O	to 1 L	
MD1	Casitone	3 g	7.6
	CaCl <sub>2</sub> • 2H <sub>2</sub> O	0.7 g	
	MgSO <sub>4</sub> • 7H <sub>2</sub> O	2 g	
	H <sub>2</sub> O	to 1 L	
TPM	Tris-HCl pH 8.0	10 mM	7.6
	MgSO <sub>4</sub>	0.8 mM	
	K <sub>2</sub> HPO <sub>4</sub> (pH 7.6)	1 mM	
	H <sub>2</sub> O	to 1 L	
MC7 (50x)	MOPS	10.5 g	7.0
	CaCl <sub>2</sub>	0.735 g	
	H <sub>2</sub> O	to 100 mL	

## 2.5 Antibiotics

**Table 2.7:** Antibiotics used in this work

Antibiotic	Stock solution	Final Concentration
Kanamycin	50 mg mL <sup>-1</sup>	50 µg mL <sup>-1</sup>
Zeocin <sup>TM</sup>	100 mg mL <sup>-1</sup>	20 µg mL <sup>-1</sup>
Ampicillin	100 mg mL <sup>-1</sup>	100 µg mL <sup>-1</sup>



## 2.6 Instruments

Table 2.8: Instruments used in this work

Instrument	Manufacturer
<b>Agarose gel electrophoresis</b>	
Electrophoresis chamber SUB-CELL®	Biorad
GT	Biorad
Electrophoresis chamber MINI-SUB-CELL® GT	
<b>Centrifuges</b>	
Cool centrifuge 5805R	Eppendorf
Cool centrifuge	Eppendorf
Tabletop centrifuge 5415D	Eppendorf
Tabletop centrifuge CT15E	Hitachi Koki
Tabletop centrifuge CT15RE	Hitachi Koki
Centrifuge Avanti JE	Beckmann Coulter
<b>Electroporator</b>	
Gene pulser XCell	Biorad
Electroporation cuvette 0.1 cm	Biorad
<b>HPLC-DAD (semi-preparative)</b>	
Agilent 1260 Infinity HPLC system	Agilent
<b>HPLC-MS-DAD (analytical)</b>	
Agilent 1100 series HPLC system	Agilent
Bruker HCTplus	Bruker
Dionex Ultimate 3000 RSLC	Dionex
LTQ-Orbitrap	Thermo Finnigan
<b>Image documentation</b>	
N-1000 Darkroom	Peqlab
Camera biovision	Peqlab
<b>Incubators</b>	
Incubators	Binder
Multitron orbital shakers	Infors
<b>Laminar flow</b>	
HeraeusLaminAir®	Kendro
Hera Safe	Kendro
<b>NMR</b>	
Bruker Avance 500	Bruker
Bruker Avance 700	Bruker

## 2 Materials and Methods

<b>PCR</b>	
Mastercycler Gradient	Eppendorf
<b>pH-measurements</b>	
pH-Meter 766 Calimatic	Knick
<b>Photometer</b>	
BioPhotometer	Eppendorf
NanoVue Plus	GE Healthcare Life Sciences
<b>Thermomixer</b>	
Thermomixer compact	Eppendorf
Thermomixer comfort	Eppendorf
<b>Sterilisation</b>	
Systec VX-150	Systec
<b>Ultrasonic bath</b>	
Sonarex	Bandelin
Vibra-Cell™	Sonics
<b>Water processing</b>	
Milli-Q water purification system	Millipore
PURELAB <i>ultra</i>	ELGA
<b>Fermenter</b>	
Labfors 3	Infors HT
Biostat B Plus	Sartorius

## 2.7 Bacterial strains, oligonucleotides, and plasmids

### 2.7.1 Bacterial strains

**Table 2.9:** Bacterial strains used in this work

Strain	Characteristics	References
<i>E. coli</i> DH10B	<i>F-mcrAΔ(mrr-hdsRMS-mcrBC)</i> <i>Ø80dlacZΔM15ΔlacX74 deoR</i> <i>recA1 endA1 araD 139Δ</i> <i>(ara,leu) 7696 galU galK rpsL</i> <i>nupG</i>	[1]
<i>M. xanthus</i> DK1622	Wildtype	[2]
<i>M. xanthus</i> DK897	Wildtype	[2]
<i>M. xanthus</i> A2	Wildtype	[3]
<i>M. xanthus</i> CSH6	Wildtype	[4]

## 2.7.2 Oligonucleotides

### 2.7.2.1 General primers used in this work

Table 2.10: General primers used in genotypic verification of mutants

Primer Label	Sequence (5' → 3')	Purification
TOPO in	CCTCTAGATGCATGCTCGAGC	DST
TOPO out	TTGGTACCGAGCTCGGATCC	DST
zeorev	AGTCCTGCTCCTCGGCCAC	DST

### 2.7.2.1 Primers used for the althiomycin project

Table 2.11: Primers used in the althiomycin project

Primer Label	Sequence (5' → 3')	Purification
Alt1A1for	CAAGGCAGGTATCTGACCAG	DST
Alt1A1rev	GAGCTCTGGCGTATGTCATC	DST
AltTE.1f	TGGAATGGAGATCTCCTGACACG	DST
AltTE.1r	CCTATCACATGGCCGGCATC	DST
AltMT.1f	CCGGTAGAGCGACACCAGG	DST
AltMT.1r	CGAGCTGCTCCATCAACTGAAG	DST
almD.1f	GTTCTCCCAGGTGTACTTCATCAG	DST
almD.1r	CAGGAAGCCCATCGACTTC	DST
Alt2Elfor	GTGATACGGAACCTCAGTCTTG	DST
Alt2Elrev	GAGAAATCTGGTACGAGGTGC	DST
Alt1A1ctrl3	GAGGACCTGGGAGCGAACTC	DST
Alt1A1ctrl4	GACCTTGCCGTTTIGATTCA	DST
AltTE.2f	CATGGTGTTTCATGCTGAGGTG	DST
AltTE.2r	CCGAACCACTCGTTCTGGTG	DST
AltMT.3f	CAGGAAGCTCATGAGCAGGTAG	DST
AltMT.3r	CGCTACCTCAACATCGAATCC	DST
almD.2f	GGAAGGTCCGGTGTGATGTC	DST
almD.2r	AGGCCAGCATCAGTCGAAG	DST
Alt2EL.2f	GGAATGGATTGAGCACGTTG	DST
Alt2EL.2r	GCACGCACGACTTCCAGG	DST
almC.7f	CTTGCGAGCAACAACGATG	DST
almC.7r	CCCAGTTCAACGACCCTGTC	DST
almC.8f	GGGACCCTCATATGACCCAGATCCTTTCGCATC	DST
almC.8r	GGAAGGAATTCTTACTTGCGAGCAACAACGATG	DST
almD.4f	CCGCGGAATTCTCAAACCTGCCAACCGATG	DST

## 2 Materials and Methods

almD.4r	AGGAGCACCATATGAAGGAACCCGCATGTAC	DST
AltDRTNde.1f	ATGAGCAACATATGCTGCTCCGCGGCAAG	DST
AltDRTEco.2r	GGCTGGAATTCTCAGTTGGCGGTGTTTCGCG	DST

### 2.7.2.2 Primers used for the inactivation of *M. xanthus* DK1622 biosynthetic gene clusters

**Table 2.12 :** Primer list for the amplification of the ~600 bp fragment of the target gene (NRPS- / PKS- encoding genes)

Target Gene	Primer Label	Sequence (5'-3')	Purification
MXAN_1291	4445.1f	CGCTTCGACCTCAGGGACGGCGT	DST
	4445.1r	GGGTCCTCAGGACGATGGTGTGAC	DST
MXAN_1600	1600.1f	GATTGCGCAGCCAGTCCTC	DST
	1600.1r	GACAACGTCCTCATCACCCTG	DST
MXAN_1601	1601.1f	GCTCCATGAGGCGTGTCACG	DST
	1601.1r	GGGGAAGAGGTGGTGCCTGAG	DST
MXAN_1602	1602.1f	GAGGACTCCGCTCACTGCTGG	DST
	1602.1r	CAGGTCGAGCACGTAGCGCTC	DST
MXAN_1603	4131.1f	ACCATCCCCTCGCCCATCGCCA	DST
	4131.1r	ACGCCGCTGGACGCGCTTGCTTC	DST
MXAN_1607	4125.1f	CGCGAATGAACCGGAGCTGATGG	DST
	4125.1r	CGGTGCTGTAGCGGGCGAACAG	DST
MXAN_2796	2966.1f	AGGGCCGCATCCTCAAGCAGC	DST
	2966.1r	AGGCGACGGAGGAAGAACGAC	DST
MXAN_2798	2964.1f	TTCAAGGGGCCCCGTCGGCGTGT	DST
	2964.1r	CGGTGATTTCATCGGGTCGCCCAC	DST
MXAN_3619	2165.1f	GCTGGGGGCGTTGAAGCTGGCT	DST
	2165.1r	GTTTCGACGCGGCCTTCTTCGG	DST
MXAN_3620	2163.1f	CTGCTGTTCAATCTGCTCTCCCGGG	DST
	2163.1r	AGCCACGCTTGTCGGTCCCCGCT	DST
MXAN_3631	2152.1f	ACCTCGACCGGATCTCCCAGC	DST
	2152.1r	CTGGATACGGAGTCGCGCTCG	DST
MXAN_3630	3630.1f	CTGCACTTCGATGGGGTCTC	DST
	3630.1r	CCTACCTGCTGCTCAACCTCATG	DST
MXAN_3634	2149.1f	GGTCTGGCGGAGACGCACGAGGT	DST
	2149.1r	GCCGCGTGGCGCTGAACGATT	DST

MXAN_3635	2148.1f	CGCAGCAGCGCGTGTTCCTCCAGT	DST
	2148.1r	GGCCTCCACGTAGCTGATGTCGTCA	DST
MXAN_3636	2147.1f	GCTGGAAACACCGTTCGACCTGTCC	DST
	2147.1r	ATGCGCGTTCGCGAAGGCCTCCT	DST
MXAN_3779	2004.1f	GAAGCGCAGTGGATGGACCCG	DST
	2001.1r	GGGTCTCCACGTAGCCGATGT	DST
MXAN_4001	1781.1f	GCGCGAGGCGGAGTGGATGGAT	DST
	1781.1r	TAGCCGATGTCGCCCCGCTCCA	DST
MXAN_4402	1380.1f	CAACGCCTGCTCCTGGAAGTCG	DST
	1380.1r	ACGGAGGGCTTGCTGGATGAC	DST
MXAN_4414	1367.1f	TCCACCTGTGGACCCGCTTGG	DST
	1367.1r	GCGGGGAAAGGGATGTCCTGATG	DST
MXAN_4532	1249.1f	AGCGTGACGCGCGGCTGAATCAGT	DST
	1249.1r	CGTGCTGGCGAAGAACCCAC	DST
MXAN_4592	1189.1f	TGGACATCGGCCAGGGGTGTC	DST
	1189.1r	GAGGATCGAGCGCGCGTTTC	DST
MXAN_4601	1180.1f	CTCCCGTTCGAGCTCCAGCAGGG	DST
	1180.1r	CTCGAGCGCCTCATGGACGGT	DST

**Table 2.13:** Control primer list for the genotypic verification of the proper integration of the knockout construct (successful disruption of NRPS- / PKS- encoding genes)

Target Gene	Primer Label	Sequence (5'-3')	Purification
MXAN_1291	4445.2f	CACGAAGCCCTGAGGACACGCT	DST
	4445.2r	ACCTCGCCCAGGAGCTGGAGCT	DST
MXAN_1600	1600.2f	CCAGGAAAGCAGAGAGCGTG	DST
	1600.2r	TGCGCACCTTCTTCATGGAG	DST
MXAN_1601	1601.2f	CCGAATACGCACGAGAGGTC	DST
	1601.2r	GCAGCAGCAGTTGGTGCAC	DST
MXAN_1602	1602.2f	CGCACAGTCACTCCACTTCTG	DST
	1602.2r	GGGAGTTGTCCGCGTAGAAG	DST
MXAN_1603	4131.2f	TGGTCCAAGGGCATCCAGCTCA	DST
	4131.2r	ACTCGTCGAGGAGTTCCTGCGT	DST
MXAN_1607	4125.2f	TGGGTGTGCTTCATGCCGCGCA	DST
	4125.2r	GCGGTTGGTGTGGGCTTCCAGC	DST
MXAN_2796	2966.2f	CTGAACCGGGAGCCTCAGCACG	DST

## 2 Materials and Methods

	2966.2r	AGGTTGGAGAGCTGCGCGGAGC	DST
MXAN_2798	2964.2f	CTTGCGGATTCGCCAGACATCG	DST
	2964.2r	TCTGCGTCTGCGTGCGGAAGGC	DST
MXAN_3619	2165.2f	AGGGACTGGTGGCGAAGTCAAT	DST
	2165.2r	GGAGCTGGAGCCTTCCGTCATG	DST
MXAN_3620	2163.2f	TCTTCCTGGAGTGTGCCTGGTCG	DST
	2163.2r	CTTCGCGGACAGCACCAGGAGC	DST
MXAN_3630	3630.2f	CCCAGGTTGGGCTTGATG	DST
	3630.2r	GGTGCTGAAGGATCCGCAC	DST
MXAN_3631	2152.2f	TTCCAGCAGTGCGTGGGCGTTGG	DST
	2152.2r	GAGCCGTGCTGGACCAGGTGGA	DST
MXAN_3634	2149.2f	ATCGACGTGGGCGGCGAGGAGCG	DST
	2149.2r	GGCGGCAGCGAAACATCACCC	DST
MXAN_3635	2148.2f	GACCACGGCTTCTTCGACATCC	DST
	2148.2r	CGAGCGGACAACGTCAACCACGT	DST
MXAN_3636	2147.2f	CCTGGACGAGGCCGTGCTGGAGCG	DST
	2147.2r	AGCATCGCCAGCGTCGCCACGGG	DST
MXAN_3779	2004.2f	CGGCTTCTTCGACCTGTCGCTGC	DST
	2001.2r	CGCCAATGCCGAAGGAGCTGAC	DST
MXAN_4001	1781.2f	AAGGTGGAGACCGCTTCGACGC	DST
	1781.2r	CGATTCCGAACGAGCTGACGCC	DST
MXAN_4402	1380.2f	CCTCGAAGACGTGCGCGCGTTC	DST
	1380.2r	AGGCTCGCTCAGGATGACGTGG	DST
MXAN_4414	1367.2f	ATGAGGCGTGGCTTCGGGAACG	DST
	1367.2r	GCGCAGATGCCCGAGCATTCTG	DST
MXAN_4532	1249.2f	AGGGTCAGCCGCGCCAGCTCCT	DST
	1249.2r	CCCTCAGCAACGCCTCCAGGTG	DST
MXAN_4592	1189.2f	GTTCTCGAAGACGAGCAGGCTC	DST
	1189.2r	CATGCAGGAGGGAATGCTGTTC	DST
MXAN_4601	1180.3f	CTCCGGCTGTACCGACGAAGCGAGA	DST
	1180.3r	TCGACGAGGAGGCTGAACGGCATGG	DST

**Table 2.14:** Primer list for the amplification of the ~600 bp fragment of the target gene (additional genes)

Target Gene	Primer Label	Sequence (5'-3')	Purification
MXAN_4698	4698.1f	GCGACAGCTCCTTGATGATG	DST
	4698.1r	CGTACTACCTGTCCCTCATCGAC	DST
MXAN_4699	4699.1f	CATGTAGCCCATGATGTCGTG	DST
	4699.1r	CGCAGTACTGTGGTCACTGCAC	DST
MXAN_6263	6263.1f	GCATGGTGGACCTGCTGAG	DST
	6263.1r	GGACCATCTCGTTGGGATGAG	DST
MXAN_4531	4531.1f	GGAGCAGTTCATCGAGCGTG	DST
	4531.1r	GGTGGCACAGGTACACGGTG	DST
MXAN_1289	1289.1f	CCATGCTGGTTGACAACCCTC	DST
	1289.1r	GTCGAAGAGGCGGCGATG	DST
MXAN_3880	3880.1f	CGGCTCCATCCTGCTCTACTC	DST
	3880.1r	GTGATGCCAATCAGGAAGCTC	DST
MXAN_3778	3778.1f	GCACTCGCCTCCGAGCTCTAC	DST
	3778.1r	GTTGAGGATGCGCTCGATGTAC	DST
MXAN_3780	3780.4f	CTGTCGAACCTGCTGCTCAC	DST
	3780.4f	ATTGGTGGCCTGACGCAG	DST
MXAN_3781	3781.1f	CGGAGGACATCTCCCTCATC	DST
	3781.1r	GAAGTGGAAGCCGAGCCAG	DST

## 2 Materials and Methods

**Table 2.15:** Control primer list for the genotypic verification of the proper integration of the knockout construct (additional genes)

Target Gene	Primer Label	Sequence (5'-3')	Purification
MXAN_4698	4698.2f	CCGTACTCGACGGCGTAATC	DST
	4698.2r	AGGAGACGTCCGCACTGTTC	DST
MXAN_4699	4699.2f	GACGAAGGGCACGTCACAG	DST
	4699.2r	TCTTCACGAGCAGGACATGTG	DST
MXAN_6263	6263.2f	ACGACACCTGGACGAACTGAC	DST
	6263.2r	GTCGTTGATGGTGCGGATG	DST
MXAN_4531	4531.2f	GCTTTCCCCAGGCGAAGT	DST
	4531.2r	CGTTTCGTTCAAGCGCCTG	DST
MXAN_1289	1289.2f	GGAACATGCACCTGCCATC	DST
	1289.2r	CCGCGTAGTGGTCAGGGAG	DST
MXAN_3880	3880.2f	GACGCCTGTCGGTGATGTT	DST
	3880.2r	CGGTGGACACCGACATCTC	DST
MXAN_3778	3778.2f	CGGCTGTCGTTCACTACGTC	DST
	3778.2r	GCCCAGGTCATAGACGAGCAC	DST
MXAN_3780	3780.2f	GTGAGAAGGTCGTCGACCTGTC	DST
	3780.2f	GCTGGTCGTCCAGGTGATG	DST
MXAN_3781	3781.2f	GCGTCTTCGCCTATGCCAC	DST
	3781.2r	TGCAGCTCGTGTGAGCTGTC	DST



## 2.7.3 Plasmids

## 2.7.3.1 General plasmids

Table 2.16: General plasmids used in this work

Plasmids	Relevant Characteristics	Reference
pCR®2.1-TOPO®	Cloning vector; <i>lacZα</i> ,	Invitrogen
pCR®II-TOPO®	T7 promoter, f1 origin, Km <sup>R</sup> , <i>bla</i> , pUC origin	
pJET1.2	Cloning vector; <i>eco47IR</i> , P <sub>lacUV5</sub> promoter, T7 promoter, <i>bla</i> , pMB1 replicon,	Fermentas
pCKT7A1	Cloning vector; T7A1 promoter, f1, origin, <i>bla</i> , <i>Sh ble</i> , Km <sup>R</sup> , pUC origin	unpublished
pCKT7A1_attb	Cloning vector; T7A1 promoter, integrase encoding gene, <i>xis</i> , attb site, <i>bla</i> , <i>Sh ble</i> , Km <sup>R</sup> , pUC origin	[5]
pDKzeo1	Cloning vector; T7A1 promoter, <i>xis</i> , integrase encoding gene, attb site, <i>Sh</i> <i>ble</i> , pUC origin	[6]
pSBTn5-kan	Cloning vector; Tn5 promoter, T7 promoter, f1 origin, Km <sup>R</sup> , <i>bla</i> , pUC origin	unpublished

## 2 Materials and Methods

### 2.7.3.2 Plasmids used for the althiomycin project

**Table 2.17:** Plasmids used in the althiomycin project

Plasmids	Target / Relevant characteristics
pAlt1A1	<i>almA</i> / pCR-TOPO-based
pAltTE	<i>almB</i> / pCR-TOPO-based
pAltMT	<i>almC</i> / pCR-TOPO-based
pTOPO-almD	<i>almD</i> / pCR-TOPO-based
pAlt2EL	<i>almF</i> / pCR-TOPO-based
pCKT7A1 <i>attb</i> AltDRT	Native <i>almE</i> cloned into pCKT7A1 <i>attb</i> for heterologous expression in <i>M. xanthus</i> DK1622
pDKzeo1-almC	Native <i>almC</i> cloned into pDKzeo1 for complementation of knockout mutant AltMT.1
pDKzeo1-almD	Native <i>almD</i> cloned into pDKzeo1 for complementation of knockout mutant almD.1

### 2.7.3.2 Plasmids used for the inactivation of *M. xanthus* DK1622 biosynthetic gene clusters

All plasmids are pCR-TOPO-based.

**Table 2.18:** Plasmids used for the targeted gene inactivation of DK1622 biosynthetic gene clusters

Plasmids	Target
pTOPO-1180.3	MXAN_4601
pTOPO-1189.1	MXAN_4592
pTOPO-1249.1	MXAN_4532
pTOPO-1367.1	MXAN_4414
pTOPO-1380.1	MXAN_4402
pTOPO-1781.1	MXAN_4001
pTOPO-2004.1	MXAN_3779
pTOPO-2147.1	MXAN_3636
pTOPO-2148.1	MXAN_3635
pTOPO-2149.1	MXAN_3634
pTOPO-3630.1	MXAN_3630
pTOPO-2152.1	MXAN_3631
pTOPO-2163.1	MXAN_3620
pTOPO-2165.1	MXAN_3619
pTOPO-2964.1	MXAN_2798
pTOPO-2966.1	MXAN_2796
pTOPO-4125.1	MXAN_1607
pTOPO-4131.1	MXAN_1603
pTOPO-1602.1	MXAN_1602
pTOPO-1601.1	MXAN_1601
pTOPO-1600.1	MXAN_1600
pTOPO-1608.1	MXAN_1608
pTOPO-4445.1	MXAN_4445

## 2 Materials and Methods

### 2.7.3.3 Plasmids used for the myxoprincomide project

**Table 2.19:** Plasmids used for the myxoprincomide project

Plasmids	Target / Relevant characteristics
pCKT7A1-A2.c506	First 650 bp of the MXAN_3779 A2 homolog cloned into pCKT7A1, Km <sup>R</sup>
pTOPOTn5MXAN3779	First 650 bp of MXAN_3779 from DK1622 was placed downstream the Tn5 promoter by assembly PCR and the final product was cloned into pCRII-TOPO
pTOPO-3778.1	MXAN_3778 / pCR-TOPO based
pTOPO-3780.4	MXAN_3780 / pCR-TOPO based
pTOPO-3781.1	MXAN_3781 / pCR-TOPO based
pTOPO-4699.1	MXAN_4699 / pCR-TOPO based
pTOPO-4698.1	MXAN_4698 / pCR-TOPO based

## 2.8 Myxococcus xanthus

### 2.8.1 Mutants of *M. xanthus* DK1622

#### 2.8.1.1 Targeted gene inactivation of unassigned biosynthetic gene clusters

**Table 2.20:** Mutants generated from the targeted gene inactivation of unassigned biosynthetic gene clusters

Strain	Targeted gene / Relevant Characteristics
1180.1	MXAN_4601 / Km <sup>R</sup>
1189.1	MXAN_4592 / Km <sup>R</sup>
1249.1	MXAN_4532 / Km <sup>R</sup>
1367.1	MXAN_4414 / Km <sup>R</sup>
1380.1	MXAN_4402 / Km <sup>R</sup>
1781.1	MXAN_4001 / Km <sup>R</sup>
2004.1	MXAN_3779 / Km <sup>R</sup>
2147.1	MXAN_3636 / Km <sup>R</sup>
2148.1	MXAN_3635 / Km <sup>R</sup>
2149.1	MXAN_3634 / Km <sup>R</sup>
3630.1	MXAN_3630 / Km <sup>R</sup>
2152.1	MXAN_3631 / Km <sup>R</sup>
2163.1	MXAN_3620 / Km <sup>R</sup>
2165.1	MXAN_3619 / Km <sup>R</sup>
2964.1	MXAN_2798 / Km <sup>R</sup>

2966.1	MXAN_2796 / Km <sup>R</sup>
4125.1	MXAN_1607 / Km <sup>R</sup>
4131.1	MXAN_1603 / Km <sup>R</sup>
1602.1	MXAN_1602 / Km <sup>R</sup>
1601.1	MXAN_1601 / Km <sup>R</sup>
1600.1	MXAN_1600 / Km <sup>R</sup>
1608.1	MXAN_1608 / Km <sup>R</sup>
4445.1	MXAN_4445 / Km <sup>R</sup>

### 2.8.1.2 Myxoprincomide-related mutants

**Table 2.21:** Mutants generated for the optimized production and biosynthetic studies of myxoprincomide

Strain	Relevant characteristics
Tn53779	Optimized DK1622 myxoprincomide producer / created using pTOPO-Tn53779 plasmid, / Km <sup>R</sup>
4699.1	Knockout mutant of MXAN_3779
4699.3	Complementation mutant of 4699.1 / Km <sup>R</sup> , <i>Sh ble</i>
3778.1	Knockout mutant of MXAN_3778 / Km <sup>R</sup>
3780.4	Knockout mutant of MXAN_3780 / Km <sup>R</sup>
3781.1	Knockout mutant of MXAN_3781 / Km <sup>R</sup>

### 2.8.1.3 Althiomycin

**Table 2.22:** Mutant of DK1622 generated for the althiomycin project

Strain	Relevant characteristics
<i>attb_almE</i>	Integration of <i>almE</i> at the <i>attB</i> site for heterologous expression, Km <sup>R</sup> , <i>Sh ble</i>

## 2 Materials and Methods

### 2.8.2 Mutants of *M. xanthus* DK897

**Table 2.23:** Mutants of *M. xanthus* DK897 generated in this work

Strain	Relevant characteristics
DK897::Alt1A1	<i>almA</i> knockout mutant, Km <sup>R</sup>
DK897::Alt2TE	<i>almB</i> knockout mutant, Km <sup>R</sup>
DK897::AltMT	<i>almC</i> knockout mutant, Km <sup>R</sup>
DK897::almD	<i>almD</i> knockout mutant, Km <sup>R</sup>
DK897::Alt2EL	<i>almF</i> knockout mutant, Km <sup>R</sup>
Tn5stopalmC	<i>almC</i> knockout mutant with Tn5 promoter insertion, Km <sup>R</sup>
Tn5stopalmD	<i>almD</i> knockout mutant with Tn5 promoter insertion, Km <sup>R</sup>
Tn5stopalmC::attb_almC	Complementation mutant of Tn5stopalmC / Km <sup>R</sup> , <i>Sh ble</i>
Tn5stopalmD::attb_almD	Complementation mutant of Tn5stopalmD / Km <sup>R</sup> , <i>Sh ble</i>

### 2.8.3 Mutants of *M. xanthus* A2

**Table 2.24 :** Mutants of *M. xanthus* A2 generated in this work

Strain	Relevant characteristics
A2.c506	Optimized A2 myxoprincomide producer / created using pCKT7A1-A2.c506, / Km <sup>R</sup>

## 2.9 Cultivation and conservation of strains

### 2.9.1 Cultivation of *E. coli*

Single colonies of *E. coli* strains were picked, inoculated into LB medium (1.5 mL), and cultivated in 2 mL Eppendorf tubes overnight (37 °C, 950 rpm). Stock cultures were prepared by mixing liquid culture (0.75 mL) with 50% v/v glycerol (0.75 ml) and stored at -80 °C. Strains were revived by inoculating fresh LB medium with a few cells from the stock culture using an inoculating loop. Media were supplemented with the respective antibiotic where appropriate.

### 2.9.2 Cultivation of *M. xanthus*

*M. xanthus* strains were grown and maintained in CTT medium. To exclude the cultivation of tan-variants, *M. xanthus* DK1622 cells were first streaked on CTT agar plates and observed for yellow colored swarms. Once growth was observed, typically after 2 days, yellow *M. xanthus* cells were scraped from the plate and inoculated into CTT medium (50 mL). For storage at -80 °C, liquid culture (0.75 mL) was mixed with

50% v/v glycerol (0.75 mL). Cells were revived by streaking a small amount of the frozen culture on CTT agar plates using an inoculating loop. *M. xanthus* mutant strains were similarly grown in CTT and CTT agar plate with the appropriate antibiotic.

## **2.10 Isolation of prokaryotic DNA**

### **2.10.1 Isolation of genomic DNA**

Isolation of genomic DNA from *M. xanthus* using the Puregene Core Kit A was performed according to the manufacturer's instructions.

### **2.10.2 Isolation of plasmid DNA by alkaline lysis**

The cell pellet from a 1.5 ml overnight culture of *E.coli* was separated from the medium by centrifugation and resuspended in resuspension buffer P1 (250 µL). Lysis buffer P2 (250 µL) was added and the mixture was shaken gently and allowed to stand at room temperature (5 min). The lysate was neutralized by adding neutralization buffer P3 (250 µL) and mixed by inverting the tube several times. The mixture was cooled in ice for 15 minutes then centrifuged (4 °C , 13000 rpm, 15 min). The supernatant containing the plasmid DNA was transferred into 1.5 mL Eppendorf tubes containing pre-cooled isopropanol (500 µL). The tubes were inverted several times to precipitate the plasmid DNA then centrifuged (4 °C , 13000 rpm, 15 min). The supernatant was discarded and the plasmid DNA pellet was washed with 70% ethanol (500 µL), air dried, and redissolved in sterile water (50 µL).

### **2.10.3 Isolation of plasmid DNA by GeneJET™ Plasmid Miniprep Kit**

Isolation of plasmid DNA with the GeneJET™ Plasmid Miniprep kit was performed according to the manufacturer's instructions.

## **2.11 Separation and purification of DNA**

### **2.11.1 Separation of DNA by agarose gel electrophoresis**

Agarose gel (0.8% agarose in TAE buffer) with ethidium bromide (~0.25 µg mL<sup>-1</sup>) was prepared. DNA samples were mixed with 6x DNA loading dye (1 µL dye per 10 µL DNA sample) and were loaded onto the gel. DNA ladders (GeneRuler™ 100 bp Plus for PCR samples, GeneRuler™ 1 kb DNA ladder for restriction digest DNA samples,

Fermentas) were used to determine the size of the DNA samples. Electrophoresis was performed in 1x TAE buffer with a voltage of approximately 10 V per cm gel length until the xylene cyanol and bromophenol blue was visibly separated by 3-4 cm. DNA was visualized under UV light (254 nm) using a gel documentation system equipped with a UV fluorescent table.

### 2.11.1 Extraction of DNA from agarose gels

After separation by gel electrophoresis, the desired DNA fragments were excised from the gel under UV light. Extraction from the gel was carried out with the NucleoSpin® Extract Kit, according to the manufacturer's protocol.

## 2.12 Polymerase chain reaction

Polymerase chain reaction (PCR) was carried out using two different polymerases. Unless otherwise stated, PCR was carried out using the *Taq* polymerase-based protocol. Amplifications of gene fragments for complementation and promoter insertion were performed using Phusion® polymerase.

### 2.12.1 Standard PCR setup

The composition of a standard PCR setup is shown in Table 2.24. The annealing temperature was adjusted depending on the primers used.

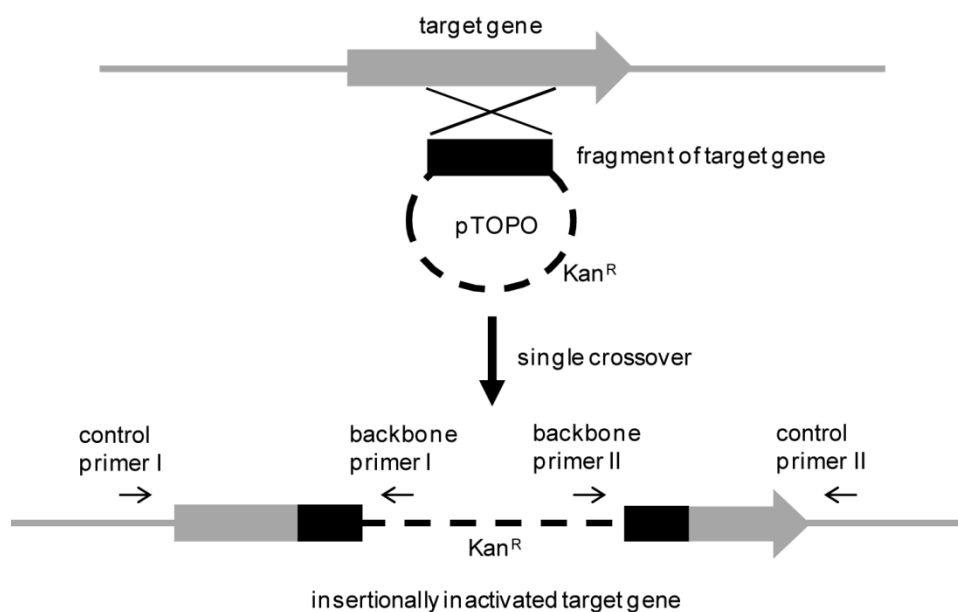


Table 2.25: PCR Protocols

	<i>Taq</i> polymerase	Phusion® polymerase
<b>Buffer</b>	10x Q buffer (2.5 µL)	5x HF buffer (5 µL)
<b>25 mM MgCl<sub>2</sub></b>	0.5 µL	-
<b>DMSO</b>	0.75 µL	1 µL
<b>Primer 1 (10 pmol µL<sup>-1</sup>)</b>	0.75 µL	0.75 µL
<b>Primer 2 (10 pmol µL<sup>-1</sup>)</b>	0.75 µL	0.75 µL
<b>25 mM dNTPs</b>	0.5 µL	0.5 µL
<b>Template</b>	0.2 µL	0.2 µL
<b>Polymerase</b>	0.2 µL	0.2 µL
<b>Sterile water</b>	18.05 µL	16.9 µL
<hr/>		
<b>Primary denaturation</b>	98 °C 3 min	98 °C, 2.5 min
<b>Denaturation</b>	98 °C 10 s	98 °C 10 s
<b>Annealing</b>	58 °C 15 s	65 °C 15 s
<b>Extension</b>	72 °C 30s / kb	72 °C 15 s / kb
<b>Final extension</b>	72 °C 8 min	72 °C 8 min
<b>Cycles</b>	30	30

### 2.12.2 PCR setup for genotypic verification of targeted gene inactivation

Recombinant *M. xanthus* strains were verified by PCR with the use of TOPO in and TOPO out primers (specific for pCR®2.1-TOPO® and pCR®II-TOPO®) and control primers specific for each inactivated gene. (Fig. 2.1).



**Figure 2.1** : Scheme of the genotypic verification of *M. xanthus* recombinant strains. Backbone primer I and backbone primer II could either be TOPO in or TOPO out primers depending on how the target gene fragment was cloned into the TOPO-based vector.

The reaction mixture is given in Table 2.25. To reduce the number of reactions, TOPO in and TOPO out primers are included in all reactions but control primers I and II are added in separate tubes. The PCR cyclyer set up is similar to that found in Table 2.24 with the annealing temperature set to 58 °C.

**Table 2.26** : Reaction mixture for the genotypic verification of recombinant *M. xanthus* strains.

	<i>Taq</i> polymerase
Buffer	10x Q buffer (2.5 µL)
25 mM MgCl <sub>2</sub>	0.5 µL
DMSO	0.75 µL
TOPO in (10 pmol µL <sup>-1</sup> )	0.75 µL
TOPO out (10 pmol µL <sup>-1</sup> )	0.75 µL
Control primer I or Control primer II	0.75 µL
25 mM dNTPs	0.5 µL
Mutant genomic DNA	0.2 µL
<i>Taq</i> Polymerase	0.2 µL
Sterile water	17.3 µL

## 2.13 Enzymatic manipulation of DNA

### 2.13.1 Restriction endonucleases

Digestions of DNA by restriction endonucleases were performed at 10  $\mu$ L scale. A typical reaction consisted of DNA (5  $\mu$ L), 10x buffer (1  $\mu$ L), sterile water (3.8  $\mu$ L), and the restriction enzyme (0.2  $\mu$ L). The restriction mixtures were incubated at 37 °C for 1-2 hours. For double digestion of DNA, the amounts of buffer and sterile water in the reaction mix were adjusted according to the manufacturer's recommendations (Fermentas).

### 2.13.2 Dephosphorylation

For the dephosphorylation of digested vector DNA, shrimp alkaline phosphatase (SAP) (1  $\mu$ L), was added to the restriction mixture. Incubation was carried out at 37 °C for 10 minutes. The enzyme was inactivated by heating the reaction at 65 °C for 15 minutes.

### 2.13.3 Standard ligation

Cloning of gene fragments into vectors was carried out in a total reaction volume of 10  $\mu$ L. Each reaction was composed of 5  $\mu$ L gene fragment, 1  $\mu$ L cloning vector, 1  $\mu$ L 10x T4 DNA ligase buffer, 1  $\mu$ L T4 DNA ligase, and 2  $\mu$ L sterile water. The ligation mix was incubated at 16 °C overnight. The reaction was heat inactivated at 65 °C for 10 minutes. The resulting plasmid DNA was subjected to phenol / chloroform extraction (Section 2.13.5) prior to electroporation into *E. coli* DH10B.

### 2.13.4 TOPO TA cloning

PCR products amplified by *Taq* polymerase were cloned into pCR®2.1-TOPO® or pCR®II-TOPO® using the TOPO TA cloning kit according to the manufacturer's protocols. Blunt end PCR products were incubated with *Taq* polymerase at 72 °C for 10 minutes prior to TOPO TA cloning to attach a single deoxyadenosine (A) to the 3' end of the PCR product.

### 2.13.5 pJET1.2 cloning

Phusion-amplified PCR products were cloned into pJET1.2/blunt cloning vector using the CloneJet™ PCR cloning kit according to the manufacturer's protocols. The constructs were extracted using phenol / chloroform prior to electroporation into *E. coli* DH10B.

### 2.13.6 Phenol / chloroform extraction of ligation products

The ligation products obtained from pJET cloning and standard ligations were extracted using phenol / chloroform prior to electroporation into *E. coli* DH10B. Nuclease-free water (80 µL) was added to 20 µL of ligation reaction mixture. A solution of phenol / chloroform / isoamyl alcohol (100 µL) was added and mixed by vortexing. The mixture was centrifuged and the aqueous layer was transferred into a new Eppendorf tube (1.5 mL) which contained 100% isopropanol (700 µL) to precipitate the plasmid DNA. The plasmid DNA was obtained as a small pellets by centrifugation. The isopropanol phase was removed and the DNA was washed with 70% ethanol (700 µL). The supernatant was removed and the DNA was dissolved in nuclease-free water (10 µL). One microliter was used to transform *E. coli* DH10B.

## 2.14 Transformation of bacteria

### 2.14.1 Electroporation of *E. coli*

Electrocompetent *E. coli* DH10B stocks were prepared by growing the strain in LB medium until OD<sub>600</sub> = 0.6-0.8 per ml. Cells were separated from the medium and washed first with 200 ml then with 20 mL of cold 10% glycerol. Finally the cells were resuspended in cold 10% glycerol (2.5 mL). The resuspended cells were distributed into 50 µL aliquots which were flash frozen in liquid nitrogen and then stored at -80 °C. To transform *E. coli*, a 50 µL aliquot was electroporated with plasmid DNA (1 µL) in an ice cold 0.1 cm cuvette (1250 V, 200 Ω, 25 mF). To transfer the cells, LB medium (1 mL) was pipetted up and down the cuvette. The cells were allowed to recover in 2 mL Eppendorf tubes (37 °C, 950 rpm, 1 h) prior to plating on LB agar plates supplemented with the appropriate antibiotic for selection of positive transformants. For *E. coli* transformed with a TOPO-based plasmid, the plates were infused with X-gal solution (40 µL, 40 mg mL<sup>-1</sup>) prior to plating.

### 2.14.2 Electroporation of *M. xanthus* strains

The *M. xanthus* strain was grown in CTT medium (50 mL) to a cell density of  $1.5 \times 10^8$  cells per mL ( $OD_{600} = 0.5$  per mL). The cells were separated from the medium by centrifugation at room temperature and washed with sterile water (2x, once with 20 mL followed by a second washing step with 1 mL). Cells were resuspended in sterile water (500  $\mu$ L). The plasmid DNA (1  $\mu$ g) was mixed with a 50  $\mu$ L aliquot of the competent *M. xanthus* cells and electroporated (650 V, 400  $\Omega$ , 25 mF) at room temperature. The cells were allowed to grow in antibiotic-free CTT medium (1 mL) in 2mL Eppendorf tubes (30 °C, 1050 rpm, 4-6 hours). After recovery, the cells were mixed with CTT soft agar equilibrated to 42 °C (5 mL) by vortexing and poured onto a CTT agar plate which was supplemented with the appropriate antibiotic. Plates were incubated at 30 °C and transformants could typically be observed 5-7 days after electroporation. Clones were picked and restreaked onto fresh CTT selection plates.

## 2.15 Genetic manipulation of *M. xanthus*

### 2.15.1 Targeted gene inactivation

Mutant strains were constructed by amplifying an internal 600 bp fragment of each NRPS or PKS encoding gene using *Taq* polymerase. The PCR product was cloned into pCR®2.1-TOPO® or pCR®II-TOPO®. The plasmid was electroporated into *M. xanthus* DK1622 (650 V / 400  $\Omega$  / 25  $\mu$ F). Transformants were selected on CTT agar plates supplemented with kanamycin (50  $\mu$ g mL<sup>-1</sup>). Proper integration of the construct into the target gene was genotypically verified using the protocol outlined in Section 2.12.2.

### 2.15.2 Complementation of inactivated genes in *M. xanthus*

The respective native gene was amplified by Phusion polymerase using primers containing an *Nde*I restriction site in the forward primer and an *Eco*RI restriction site in the reverse primer. The PCR product was digested using the aforementioned restriction endonucleases and cloned into pDKzeo1. The resulting construct was electroporated into the corresponding knockout mutant *M. xanthus* strain and plated on a CTT agar plate supplemented with kanamycin (50  $\mu$ g mL<sup>-1</sup>) and zeocin (20  $\mu$ g mL<sup>-1</sup>). Genotypic verification by PCR was performed in a first step to verify the original inactivation by use of the control primers and in a second step the integration of the copy intended for complementation was verified using zeorev primer and a primer internal of the

respective gene. The phenotype was determined by growing the wildtype, the knockout mutant, and the complementation mutant in production cultures from which extracts were prepared that were analyzed by HPLC-HRMS.

### 2.15.3 Tn5 promoter insertion by single crossover homologous recombination

A ~600 bp internal fragment of the target gene was amplified using Phusion polymerase. The forward primer contained an in-frame substitution of a codon with a stop codon. The Tn5 promoter region was amplified separately. The final products were assembled by PCR using *Taq* polymerase and cloned into pCR2.1-TOPO. Alternatively, the ~600 bp internal fragment was amplified using primers with *NdeI* and *EcoRI* restriction sites. The PCR product was digested with *NdeI* and *EcoRI* and cloned into pSBTn5. Final constructs were transformed into *M. xanthus*.

### 2.15.4 Heterologous expression of *almE* in *M. xanthus* DK1622

The gene *almE* was amplified using AltDRTNde.1f and AltDRTEco.2r primers. The PCR product was digested with *NdeI* and *EcoRI* restriction enzymes and cloned into pCK-T7A1 *attB*.<sup>[5]</sup> The plasmid was electroporated into *M. xanthus* DK1622 and the transformants were plated on CTT agar plates supplemented with zeocin (20 µg mL<sup>-1</sup>). Mutants were verified by PCR using primers binding to *almE*. The phenotype was determined by disc assays against wildtype and the DK1622 mutant harboring the *almE* gene to display different althiomycin-related growth inhibition zones.

## 2.16 Analysis of secondary metabolite production of *M. xanthus* strains

### 2.16.1 Preparation of production cultures

Wildtype and recombinant *M. xanthus* strains were grown in CTT medium (50 mL) supplemented with Amberlite XAD-16 resin (2%, Sigma-Aldrich) and the appropriate antibiotic (for mutant strains). Cultures were grown in Erlenmeyer flasks (250 mL) at 30 °C for 96 h at 200 rpm. Cell and XAD resin pellets were harvested by centrifugation (20 min, 5000 rpm, 20 °C). Pellets were extracted with methanol (2 x 30 mL). Extracts were dried *in vacuo* and resuspended in distilled methanol (1 mL).

### 2.16.2 LC-HRMS analysis

LC-HRMS was carried out with a Dionex Ultimate 3000 RS LC coupled to a LTQ Orbitrap (Thermo Scientific) equipped with a TriVersa NanoMate (Advion) ion source. Separation was performed on an RP-18 column (50 × 2.1 mm, particle size 1.7 mm, Waters Acquity BEH) with a mobile-phase system consisting of water and acetonitrile (each containing 0.1% formic acid) using a gradient from 5 – 95% of the organic phase in 12 minutes. Data sets were acquired in positive electrospray (ESI) mode with a scan range from 200–2000 *m/z*.

## 2.17 Sequence analysis of the *Myxococcus xanthus* PKS / NRPS biosynthetic gene clusters

Routine DNA and protein *in silico* sequence analysis tasks were carried out using the Geneious v5 Software.<sup>[7]</sup>

### 2.17.1 Annotation of the *M. xanthus* DK1622 biosynthetic gene clusters

The complete genome sequence of *M. xanthus* DK1622 is publicly available (Genbank accession no. NC\_008095). Sequences of the unassigned biosynthetic gene clusters were extracted and initial annotation of the catalytic domains was performed using the PKS/NRPS analysis software (<http://nrps.igs.umaryland.edu/nrps/index.html>).<sup>[8]</sup> Additional prediction of A domain substrates was done using the NRPSpredictor2 software.<sup>[9]</sup>

### 2.17.2 Annotation of the althiomycin biosynthetic gene cluster from *M. xanthus* DK897

Open reading frames were assigned using FramePlot web software (<http://nocardia.nih.go.jp/fp4/>). Annotation of the NRPS and PKS catalytic domains and the prediction of A domain substrates were performed using the same protocol outlined in Section 2.16.1. The gene cluster was deposited in the EMBL/Genbank database with the accession number FR831800.

### 2.17.3 Domain alignment and phylogenetic analysis

All sequence alignments were performed by ClustalW.<sup>[10]</sup> Phylogenetic trees were generated using the Mr. Bayes plugin of the Geneious software suite. The Bayesian estimation for the phylogenetic analysis of TE domains used the same parameters used by Buntin, *et al.*<sup>[11]</sup>

### 2.18 Comprehensive analysis of *M. xanthus* DK1622 secondary metabolome

#### 2.18.1 Preparation of *M. xanthus* sample sets

*M. xanthus* wild-type and mutant strains were grown in replicates (n=4) in 50-mL scale cultures and the corresponding extracts were prepared as described in Section 2.15.1.

#### 2.18.2 Liquid chromatography – high resolution mass spectrometry (LC-HR-MS)

LC-HR-MS method development was carried out using an UltiMate 3000™ rapid separation liquid chromatography system (Dionex RSLC) coupled to an UHR-TOF mass spectrometer (Bruker Daltonik maXis) operating in positive ESI mode. LC-nanoESI measurements were performed using an Accela UPLC system coupled to an LTQ Orbitrap mass spectrometer (Thermo Scientific) equipped with a TriVersa NanoMate nanoESI source (Advion). Separation was performed on an RP-18 column (50 mm x 2.1 mm, particle size 1.7  $\mu$ m; BEH, Waters) using a mobile phase system consisting of water and acetonitrile (each containing 0.1% formic acid). The following gradient was applied for separation with a flow rate of 400  $\mu$ L/min using (A) Water + 0.1% formic acid (B) Acetonitrile + 0.1% formic acid as mobile phase: 0 min 1% B; 1 min 1% B; 10 min 99% B; 12 min 99% B; 12.5 min 1% B; 14 min 1% B. Data sets were acquired with a scan range from 200-2000  $m/z$ . Targeted verification was performed by creating individual extracted ion chromatograms of the candidate  $m/z$  values generated by PCA and *t*-test models.

#### 2.18.3 Statistical treatment of data

PCA and *t*-test evaluation were both performed using the ProfileAnalysis v. 2.0 software (Bruker Daltonik GmbH) using the find molecular features (FMF) tool. For each sample set four replicates (n=4) were used. The LC-MS data were processed from 2 to 11 minutes and mass ranges of 200 – 1000  $m/z$  and 1000 -2000  $m/z$ . The bucketing parameters for the FMF approach were  $\Delta m/z$  = 30 mDa and  $\Delta RT$  = 30 s. Compound detection was performed using the following parameters : S/N = 5; Correlation coefficient threshold = 0.7; Minimum compound length = 10 spectra; Smoothing width = 2. To improve visualization of clusters in the scores and loadings plot, the bucket table was subjected to pretreatment by log transformation and pareto scaling.<sup>[12]</sup> Pairwise comparison of wild-type and mutant strain sample sets were performed for both the PCA and *t*-test evaluation.



## 2.19 Yield optimization and large scale fermentation of myxoprincomide

### 2.19.1 Determination of the best producer

To determine the best producer for myxoprincomide, five myxoprincomide-producing *M. xanthus* strains (A23, A2, A5, A85, Mxx17) were selected and grown quantitatively against DK1622 at 15-ml scale (14.4 mL CTT, 0.3 mL XAD-16, 0.3 mL inoculum) in replicates. All 5 strains were grown to an  $OD_{600} / \text{ml} = 0.5$  prior to inoculation to ensure an equal starting point. The XAD and cell pellet were extracted with methanol (30 mL) and concentrated to final volume of 300  $\mu\text{L}$ . The corresponding extracts were analyzed by HPLC-MS (Section 2.18.3).

### 2.19.2 Media optimization

*M. xanthus* DK1622 and A2 were grown in 15-ml scale production cultures in 4 different media: CTT, CYE, MD1, and TS. After extraction, the samples were analyzed by HPLC-MS (Section 2.18.3)

To significantly reduce the production of myxochelin, which coelutes with myxoprincomide, TS medium was supplemented with Fe(III)-EDTA (80  $\mu\text{g mL}^{-1}$ ).

Further comparison of myxoprincomide production was conducted between TS and VY/2 medium. Similar to previous experiments, the cultures were grown at 15-mL scale and the extracts were prepared. The extracts were analyzed by LC-HR-MS (Section 2.17.2). Peak integration of myxoprincomide ( $[M+2H]^{2+} = 506.2717 \text{ m/z}$ ) was performed using the Xcalibur QualBrowser software.

### 2.19.3 Quantitative analysis of *M. xanthus* extracts

Extracts were analyzed by HPLC-MS on an Agilent 1100 series solvent delivery system coupled to Bruker HCTplus ion trap mass spectrometer. Chromatographic separation of a 5  $\mu\text{L}$  sample was carried out on a Luna C18(2) column (100  $\times$  2 mm, 2.5- $\mu\text{m}$  particle size; Phenomenex) equipped with a precolumn C18 (8  $\times$  3 mm, 5  $\mu\text{m}$ ). The mobile-phase gradient was composed of solvent A [water + 0.1% formic acid] and solvent B [acetonitrile + 0.1% formic acid] linear from 5% B at 1 min to 95% B at 12 min at a flow rate of 400  $\mu\text{L min}^{-1}$ . Quantitation was carried out in manual MS<sup>2</sup> mode. Ions of  $m/z$   $[M+2H]^{2+} = 506.3$  were collected and subjected to fragmentation. The intensities of the

## 2 Materials and Methods

---

characteristic fragment ions  $m/z = 797.5$  and  $m/z = 497.4$  were summed up, and peak integration was carried out utilizing the Bruker DataAnalysis software package.

### 2.19.4 Strain optimization

*M. xanthus* A2 was engineered to produce higher myxoprincomide titers by promoter insertion directly before the gene MXAN3779. A copy of the first 650 bp of MXAN\_3779 was linked to the T7A1 promoter sequence by cloning the fragment into the plasmid pCKT7A1. Single crossover recombination in *M. xanthus* A2 resulted in two copies of the 650 bp MXAN\_3779 fragment. The first copy of MXAN\_3779 is truncated whereas the second copy represent the in-frame and full-length gene under control of the T7A1 promoter. The first 650 bp of MXAN 3779 were amplified with Phusion polymerase using primers with *NdeI* (MXAN3779\_startNdeI - 5'-CGCTCATATGCACGAGACTCCGCGAAC) and *EcoRI* (MXAN3779\_Eco - 5'-ATGGAATTCTACGGCGGCAGCCAGATGAC) restriction sites designed using the available *M. xanthus* DK1622 genome sequence. The PCR product was cloned into the pCKT7A1 vector and sequenced. The plasmid was electroporated into *M. xanthus* A2 (650 V / 400  $\Omega$ / 25  $\mu$ F). Transformants (A2.c506) were verified by PCR amplification using a forward primer specific for pCKT7A1(5'- CACGTGCTGATCAGATCCATTAC) and a reverse primer specific for MXAN 3779 (5'- GGAGGATGCCTCCGATGAG) and grown quantitatively against *M. xanthus* A2 wild type with a starting OD<sub>600</sub> / mL= 0.25.

#### 2.19.3.2 Optimization of the introduction of XAD-16 resin

Four pairs of 15 ml production cultures of the optimized strain A2.c506 were prepared using TS medium supplemented with Fe(III)-EDTA and kanamycin. XAD-16 resin (300  $\mu$ L) was added at different time points during cultivation: Sample set A: start of cultivation, sample set B: after 24 hours, sample set C: after 48 hours, and sample set D: after 72 hours. Sample sets A to C were harvested 72 hours after inoculation. Sample set D was allowed to grow for an additional 72 hours (death phase). The extracts were prepared as previously described and the samples were analyzed by HPLC-MS (Section 2.18.3)

### 2.19.4 Large scale fermentation

To obtain sufficient amount of myxoprincomide, the optimized strain *M. xanthus* A2.c506 was grown in PS medium (300-L) and VY/2 (fed-batch 9-L)

#### 2.19.4.1 300-L PS medium

The optimized myxoprincomide producer strain *M. xanthus* A2.c506 was grown at 300-L scale in PS medium (10 g Casein Peptone Type S (Marcor) 2 g MgSO<sub>4</sub> 80 µg/ml Fe(III) EDTA and 4 g starch per liter pH 7.0) for 72 hours before the addition of 2% XAD-16 resin. Fermentation was carried out at 250 rpm, 30 °C, and pH was maintained at 7.0. After XAD introduction, the cultivation was allowed to proceed for an additional 72 hours before harvesting. The cells and XAD-16 resin were separated from the medium and amounted to 6.6 kg.

#### 2.19.4.2 9-L fed-batch VY/2 fermentation

Six liters of VY/2 medium (without HEPES) supplemented with kanamycin (50 µg mL<sup>-1</sup>), vitamin solution (100 µL per liter medium) and trace elements (2 mL per liter medium) were inoculated with 500 mL of *M. xanthus* A2.c506 grown in VY/2 + Kan50. Fermentation was run at 250 rpm, 30 °C, pH 7.0 and the pO<sub>2</sub> was maintained by flow. The feed solution 10x VY/2 (50 g Baker's yeast, 0.5 g CaCl<sub>2</sub>, H<sub>2</sub>O to 1 L) and XAD-16 were added once the logarithmic phase had started. The feeding rate was manually adjusted to maintain a linear oxygen consumption curve. Cultivation was terminated at the onset of the stationary phase.

### 2.20 Purification of myxoprincomide from the optimized producer strain *M. xanthus* A2.c506

#### 2.20.1 Isolation of myxoprincomide-c506 from 300-L PS medium fermentation of *M. xanthus* A2.c506

The cells and XAD-16 resin from the 300-L scale fermentation were extracted with methanol. Before concentration, the methanol extract was defatted with n-heptane. The methanol phase was dried *in vacuo*. The dried residue (~30g) was washed with ethyl acetate. The excess ethyl acetate was removed and the remaining residue was partitioned between water and n-butanol (1:1). The water phase was lyophilized at -80°C. 2.5 g of the dried residue was immersed in 10 ml of methanol. The mixture was centrifuged and the supernatant was fractionated on a Sephadex LH-20 column using methanol as the eluent. Aliquots from fractions were obtained and analyzed by HPLC-MS. Fractions containing myxoprincomide were pooled, dried *in vacuo*, and re-dissolved in double distilled water (500 µl). The myxoprincomide-containing fraction

## 2 Materials and Methods

---

was purified by reverse-phase HPLC (Jupitor 4u Proteo 90A, 250 x 10 mm, 4 micron, DAD at 220 nm and 280 nm) using a H<sub>2</sub>O-acetonitrile containing 0.05% trifluoroacetic acid. Elution was carried out as follows: 15% ACN (0 - 5 min), 15 - 20% ACN (5-10 min), 20% ACN isocratic (10-20 min). Myxoprincomide (1.2 mg,  $t_R$  = 15.1 min) was collected and dried *in vacuo* at -80°C prior to NMR analysis.

### 2.20.2 Isolation of myxoprincomide variants from 9-L fed-batch VY/2 fermentation of *M. xanthus* A2.c506

The XAD-16 resin was separated from the culture broth and was washed with water. The resin was defatted with *n*-heptane followed by extraction with ethyl acetate until most of the DKxanthenes was removed from the XAD-16 beads, judged by visual inspection. Final extraction was carried out with methanol until the supernatant remained colorless. The methanol fraction was dried *in vacuo* and the dried residue was redissolved in 10 mL methanol. The concentrated sample was fractionated on a Sephadex LH-20 column with methanol as the mobile phase. Fractions were analyzed by HPLC-MS. Fractions containing myxoprincomide variants were pooled, completely dried, and redissolved in double distilled water (100 µL per 10 mg dried residue). Purification of the myxoprincomides by reverse-phase HPLC using the exact same conditions detailed in Section 2.20.1.

## 2.21 Structure elucidation of myxoprincomide

### 2.21.1 NMR Analysis

NMR spectra for the myxoprincomides were recorded on a Bruker Avance 700 instrument at 700 MHz and 175 MHz in DMSO-*d*<sub>6</sub> as a solvent and internal standard.

### 2.21.2 Determination of the absolute stereochemistry of *N*-MeSer, Ser, Leu, OH-Val, Tyr, β-Lys, and Ala

Purified myxoprincomide-c506 (0.6 mg) was hydrolyzed with 6N HCl in a sample vial for 16 h at 100 °C. The hydrolysates were evaporated to dryness, dissolved in H<sub>2</sub>O (100 µl), and aliquoted to 50 µl portions. 1N NaHCO<sub>3</sub> (20 µl) and 1% 1-fluoro-2,4-dinitrophenyl-5-L-leucinamide (L-FDLA or D-FDLA solution in acetone, 100 µl) were added to each aliquot. The mixtures were incubated at 37 °C for 1h. The solutions were cooled to room temperature, neutralized with 2N HCl (20 µl), and evaporated to dryness. The residues were dissolved in acetonitrile (810 µl) and analyzed by LC-MS.

HPLC-DAD-MS analysis of the L/D-FDLA derivatives was performed on an HPLC-DAD-MS system operated in switching ionization mode. Separation was achieved using a Luna RP-C<sub>18</sub> column (100 mm x 2 mm, particle size 2.5 µm, Phenomenex) using a mobile phase system consisting of water and acetonitrile containing 0.1 % formic acid. Elution was performed with a linear gradient of 5 – 95% acetonitrile in 20 min at a flow rate of 0.4 ml / min. MS data were acquired at a scan range of 100 – 1400 *m/z*.

### 2.21.3 Determination of the absolute stereochemistry of the 2-oxo-beta-leucine in myxoprincomide-c506

The following procedure was adapted from the method described by Itagaki *et al.*<sup>[13]</sup> To a stirred solution of myxoprincomide-c506 (0.3 mg) in 5% NaOH (0.5 ml), 35% H<sub>2</sub>O<sub>2</sub> (0.1 ml) was added dropwise to the reaction. The reaction was stirred at 65 °C for 40 min after which the reaction was hydrolyzed with 6N HCl (1.0 ml) at 100°C for 16h. The reaction mixture was dried *in vacuo* and the hydrolysates were derivatized with FDLA as previously described.

### 2.20.4 Tandem MS fragmentation

MS<sub>n</sub> analyses were carried out on a LTQ Orbitrap mass spectrometer by collision induced dissociation (CID) at normalized collision energies of 35% at a mass resolution of *R* = 30 000. Parent ions were isolated with a 2 *m/z* window. For HPLC-MS/MS<sup>2</sup> analyses, full scan measurements were acquired in the Orbitrap and fragments were measured in the LTQ. For direct infusion of purified compounds, full scan and fragment measurements were carried out in the Orbitrap.

## 2.21 Feeding of labeled substrates

### 2.21.1 CTT (20 mL)

Shake flask production cultures (20-mL) of *M. xanthus* DK1622 were supplemented with 1 mM of labelled amino acids. The following amino acids were used: L-alanine-2-d; L-leucine-5,5,5-d<sub>3</sub>; L-leucine-2,3,3,4,5,5,5',5',5'-d<sub>10</sub>; L-glutamic-2,3,3,4,4-d<sub>5</sub> acid; L-serine-2,3,3-d<sub>3</sub>; L-valine-d<sub>8</sub>; L-serine- <sup>13</sup>C<sub>3</sub>, <sup>15</sup>N; L-threonine - <sup>13</sup>C<sub>4</sub>, <sup>15</sup>N; 5-fluoro-DL-tryptophan; L-phenylalanine-<sup>13</sup>C<sub>6</sub>; L-histidine-<sup>13</sup>C<sub>6</sub>; L-methionine(methyl-d<sub>3</sub>); sodium 1-<sup>13</sup>C-acetate; sodium 2-<sup>13</sup>C-acetate; & 1,2-<sup>13</sup>C<sub>2</sub>-acetate. Cultures were grown for 96 hours at 30°C , 200 rpm. The cells and XAD were separated from the medium by centrifugation and were

extracted with 2 x 15 ml methanol. The extracts were concentrated to a final volume of 300 µl and analyzed by HPLC-HR-MS.

### 2.21.2 VY/2 (50 mL)

Shake flask production cultures (50-mL) of the *M. xanthus* strain were supplemented with 1 mM of labelled amino acids. The following amino acids were used: L-alanine-2-d; L-leucine-5,5,5-d<sub>3</sub>; L-valine-d<sub>8</sub>; L-serine- <sup>13</sup>C<sub>3</sub>, <sup>15</sup>N; L-phenylalanine-<sup>13</sup>C<sub>6</sub>; L-methionine(methyl-d<sub>3</sub>); sodium 1-<sup>13</sup>C-acetate; sodium 2-<sup>13</sup>C-acetate; & 1,2-<sup>13</sup>C<sub>2</sub>-acetate. Cultures were grown for 96 hours at 30°C , 200 rpm. The cells and XAD were separated from the medium by centrifugation and were extracted with 2 x 30 ml methanol. The extracts were concentrated to a final volume of 1 mL and analyzed by HPLC-HR-MS.

## 2.22 Isolation and structure elucidation of althiomycin

*Myxococcus xanthus* DK897 was grown in a 300 L fermenter with PS medium (10 g/L Casein Peptone Type S (Marcor), 2 g/L MgSO<sub>4</sub>, 80 µg/ml Fe(III) EDTA , and 4 g/L starch) pH = 7.09). After 72 h, 2 % XAD-16 resin was added. The culture was harvested after a total of 144 h of cultivation, and XAD and the cells were separated from the liquid. The XAD cell mixture was subsequently extracted with water, a mixture of water/methanol 1:1, and pure methanol. The methanol fraction contained althiomycin. The volume of the methanol fraction was reduced and the concentrated methanol fraction was defatted with *n*-heptane. The methanol phase was dried *in vacuo*. The crude extract was dissolved in water and extracted with ethyl acetate. Major parts of althiomycin stayed in the ethyl acetate fraction. This fraction was dried and applied to a Sephadex LH-20 column and eluted using methanol. The althiomycin containing fractions were pooled and concentrated.

A further purification step was performed via a Waters AutoPurification System with a Xbridge Prep C18 OBD column (19 x 150 mm, 5 µm particle size; Waters) using buffered water (5 mM NH<sub>4</sub>HCO<sub>2</sub> pH 5.8) as solvent A and methanol (containing 5 % of A) as solvent B. A gradient of 40 – 95% B in 6.5 min was initiated after 2 min. Detection was performed by a 3100 Mass Detector (Waters) and fractions containing althiomycin were collected on a 2767 Sample Manager (Waters), pooled, and concentrated.

NMR spectra for althiomycin were recorded on a Bruker Avance 500 instrument at 500 MHz and 125 MHz in DMSO- $d_6$  as a solvent and internal standard.

## 2.23 Bioactivity assays

### 2.23.1 Althiomycin sensitivity assay on *M. xanthus* DK1622 wildtype and *attb\_almE* mutant strains

Paper discs were soaked with 15  $\mu$ l of purified althiomycin and dehydrated althiomycin with concentrations of 1 mg ml<sup>-1</sup>. The discs were placed on top of *M. xanthus* DK1622 wild-type and *attb\_almE* lawns. Detection of inhibition zones was performed by visual inspection.

### 2.23.2 Activity screening for myxoprincomide

#### 2.23.2.1 Cell culture

Cell lines were obtained from the American Type Culture Collection (ATCC) and the German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung für Mikroorganismen und Zellkulturen, DSMZ). All cell lines were cultured under conditions recommended by the respective depositor. Media were purchased from Sigma, fetal bovine serum (FBS Gold) from PAA, and other reagents from GIBCO (Invitrogen). The stable cell line CHO/NF $\kappa$ Bp65-GFP was obtained from Affymetrix and grown in medium supplemented with hygromycin (100  $\mu$ g mL<sup>-1</sup>). U-2 OS cells expressing GR-EGFP were generated by lentiviral transduction and calcium phosphate transfection of an expression vector encoding a fusion protein of the glucocorticoid receptor and EGFP.

#### 2.23.2.2 Cytotoxic activity

Myxoprincomides were dissolved in 60  $\mu$ l cell culture medium to a concentration of 300  $\mu$ g/ml in 96-well cell culture plates. Then, 120  $\mu$ l cell suspension (5 x 10<sup>4</sup> cells/ml) of either HCT-116 (human colon carcinoma) or CHO-K1 (Chinese hamster ovary) cells were added to achieve a final drug concentration of 100  $\mu$ g/ml. Viability of cells was assessed after 4 d treatment by tetrazolium salt reduction (MTT) and compared to the solvent control.

### 2.23.2.3 ROS assay

U-2 OS cells were grown in 96-well imaging plates until they reached approximately 80% confluency and washed with KRH buffer (Krebs-Ringer-HEPES; 119 mM NaCl, 4.7 mM KCl, 2.5 mM CaCl<sub>2</sub>, 1.2 mM MgCl<sub>2</sub>, 1.2 mM KH<sub>2</sub>PO<sub>4</sub>, 10 mM HEPES, pH 7.4). Cells were loaded with 5  $\mu$ M DCFDA (100  $\mu$ l/well) for 1 h at 37°C. After washing with KRH and 30 min recovery in cell culture medium, cells were stimulated 30 nM tBHP (*tert*-butyl hydroperoxide) in order to artificially stimulate ROS (reactive oxygen species) production. In parallel, counteraction (ROS scavenging) of myxoprincomides was assessed at a final concentration of 100  $\mu$ g/ml and compared to 50  $\mu$ M Resveratrol (positive control; described antioxidant). DCFDA is oxidized by ROS and upon that shows strong fluorescence in the GFP channel. Images were recorded on an automated microscope (BD Pathway855) and processed in AttoVision v1.6.2. Fluorescence intensity was determined on a single cell basis and averaged over one well.

### 2.23.2.4 NF $\kappa$ B inhibition

CHO/NF $\kappa$ Bp65-GFP cells were grown in 96-well imaging plates in F12 medium supplemented with 10% FBS / hygromycin (100  $\mu$ g mL<sup>-1</sup>) until 80-90% confluency. Medium was changed to serum-free F12 and cells were incubated for 3 h at 37°C. Myxoprincomides were added at a final concentration of 100  $\mu$ g/ml and cells were treated for 1 h prior to the addition of 25 ng/ml IL-1 $\beta$ . By this, NF $\kappa$ Bp65 translocation from the cytoplasm to the nucleus can be observed and it is expected that anti-inflammatory compounds that act via NF $\kappa$ B signaling are able to inhibit this stimulus-induced translocation. After 30 min stimulation at 37°C cells were fixed with 4% PFA (paraformaldehyde) and nuclei were stained with Hoechst33342 (Molecular Probes). Images were recorded on an automated microscope (BD Pathway855) and processed in AttoVision v1.6.2. Hoechst images were used for segmentation on nuclei. The GFP fluorescence was determined in both, cytoplasmic and nuclear segments, as average value of cells within the field of view. After scaling, ratios of GFP fluorescence in the nuclei and cytoplasmic regions were calculated.

### 2.23.2.5 GR translocation

U-2 OS/NR3C1-GFP cells were cultured in 96-well imaging plates and treated with either 100  $\mu$ g mL<sup>-1</sup> myxoprincomides or 100 nM dexamethasone (Dex) as a positive control. Glucocorticoid receptor ligands like Dex lead to a translocation of GR from the



cytoplasm to the nucleus and by this, generally suppress immune and inflammatory responses. After 3 h treatment at 37°C cells were fixed with 4% PFA (paraformaldehyde) and nuclei were stained with Hoechst33342 (Molecular Probes). Images were recorded on an automated microscope (BD Pathway855) and processed in AttoVision v1.6.2. Hoechst images were used for segmentation on nuclei. The GFP fluorescence was determined in both, cytoplasmic and nuclear segments, as average value of cells within the field of view. After scaling, ratios of GFP fluorescence in the nuclei and cytoplasmic regions were calculated.

### 2.23.2.6 Bacterial cultures

All microorganisms were handled under standard conditions recommended by the depositor. Overnight cultures of bacteria were prepared in EBS medium (0.5% peptone casein, 0.5% proteose peptone, 0.1% peptone meat, 0.1% yeast extract; pH 7.0) and of yeast and fungi in Myc medium (1% phytone peptone, 1% glucose, 50 mM HEPES, pH 7.0) by inoculation either from cryocultures or of single colonies on agar plates. The next day, OD<sub>600</sub> was measured on a photometer.

**Table 2.27:** List of strains used for the bioactivity assays

Strain	Source (Strain collection)	Temperature [°C]
<i>Candida albicans</i>	HZI / DSMZ (DSM 1368)	30
<i>Mucor hiemalis</i>	HZI	30
<i>Pichia anomala</i>	HZI / DSMZ (DSM 70263)	30
<i>Chromobacterium violaceum</i>	HZI	30
<i>Pseudomonas syringae</i>	HIPS	30
<i>Pseudomonas aeruginosa</i>	HZI	30
<i>Escherichia coli</i> TolC	HZI / Ciba-Geigy AG, Basel	37
<i>Bacillus subtilis</i>	HIPS	30
<i>Mycobacterium smegmatis</i> mc <sup>2</sup> 155	ATCC (700084)	37
<i>Staphylococcus aureus</i> MRS3	HZI	30
<i>Mycobacterium diernhoferi</i>	HZI	30
<i>Micrococcus luteus</i>	HZI	30

### 2.23.2.7 Antimicrobial testing

Overnight cultures of microorganisms were diluted to OD<sub>600</sub> 0.01 (bacteria) or 0.05 (yeast/fungi) in the respective medium. Inhibition experiments were performed in 96-well plates at a final drug concentration of 200 µg/ml in a total volume of 150 µl suspension culture/well. After overnight treatment, growth inhibition was assessed by OD<sub>600</sub> measurement on a plate reader.

### 2.23.3 Development Assay, sporulation frequency and motility rates

The wildtype and mutant strains were grown to a cell count of  $0.6 \times 10^8$  cells per mL. One milliliter of the cultures was centrifuged for 5 minutes at 10,000 rpm at room temperature. The supernatant was discarded and the cell pellet was washed with MC7 medium. The cell suspension was centrifuged for 5 minutes at 10,000 rpm. The cell pellet was resuspended in MC7 medium (15  $\mu$ L) in order to have a final cell concentration of  $5 \times 10^9$  cells per mL. The cell suspension (5  $\mu$ L) was spotted on TPM agar in triplicates. The spots were allowed to dry. The plates were incubated at 30 °C. Photographs were taken after 6h, 12h, 24h, 48, and 72h at x10 magnification through a dissecting microscope.

Measurement of the sporulation frequencies and motility rates were conducted using the protocol described by Velicer *et al.* with a few modifications.<sup>[14]</sup>

Preparation of the starvation plates for the sporulation frequency assay was identical to the procedure used for the development assay except that 50  $\mu$ L of the cell suspension was spotted on the TPM agar.

For the motility assay, 50  $\mu$ L of culture was spotted on soft and hard CTT agar plates. Swarm perimeters were marked after 72 and 144 hours of incubation.

## 2.24 References

- [1.] S. G. Grant, J. Jessee, F. R. Bloom, D. Hanahan, *P.Natl.Acad.Sci.USA* **1990**, 87 4645-4649.
- [2.] D. Kaiser, *P.Natl.Acad.Sci.USA* **1979**, 76 5952-5956.
- [3.] M. Vos, G. J. Velicer, *FEMS Microbiol.Ecol.* **2008**, 64 343-350.
- [4.] D. Krug, G. Zurek, O. Revermann, M. Vos, G. J. Velicer, R. Müller, *Appl.Environ.Microbiol.* **2008**, 74 3058-3068.
- [5.] H. B. Bode, M. W. Ring, G. Schwär, M. O. Altmeyer, C. Kegler, I. R. Jose, M. Singer, R. Müller, *ChemBioChem* **2009**, 10 128-140.
- [6.] D. Krug, R. Müller, *ChemBioChem* **2009**, 10 741-750.
- [7.] Drummond, A. J., Ashton, B., Buxton, S., Cheung, M., Heled, J., Kears, M., Moir, R., Stones-Havas, S., Thierer, T., and Wilson, A. Geneious Pro 4.8.3 . 2010.
- [8.] B. O. Bachmann, J. Ravel, *Methods in Enzymology* **2009**, 458 181-217.
- [9.] M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch, O. Kohlbacher, *Nucleic Acids Res.* **2011**, 39 W362-W367.
- [10.] R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, J. D. Thompson, *Nucleic Acids Res.* **2003**, 31 3497-3500.
- [11.] K. Buntin, K. J. Weissman, R. Müller, *ChemBioChem* **2010**, 11 1137-1146.
- [12.] R. A. van den Berg, H. C. Hoefsloot, J. A. Westerhuis, A. K. Smilde, M. J. van der Werf, *Bmc Genomics* **2006**, 7 142.
- [13.] F. Itagaki, H. Shigemori, M. Ishibashi, T. Nakamura, T. Sasaki, J. Kobayashi, *J.Org.Chem.* **1992**, 57 5540-5542.
- [14.] G. J. Velicer, L. Kroos, R. E. Lenski, *P.Natl.Acad.Sci.USA* **1998**, 95 12376-12380.

### 3 Myxoprincomide: a natural product from *Myxococcus xanthus* discovered by comprehensive analysis of the secondary metabolome

Cortina, N. S.\*, Krug, D.\* , Plaza, A., Revermann, O., Müller, R. *Angew. Chem. Int. Edit.* **2012** 51, 811 – 816. (\*Authors contributed equally to this work)

This article is available online at:

<http://onlinelibrary.wiley.com/doi/10.1002/anie.201106305/pdf>

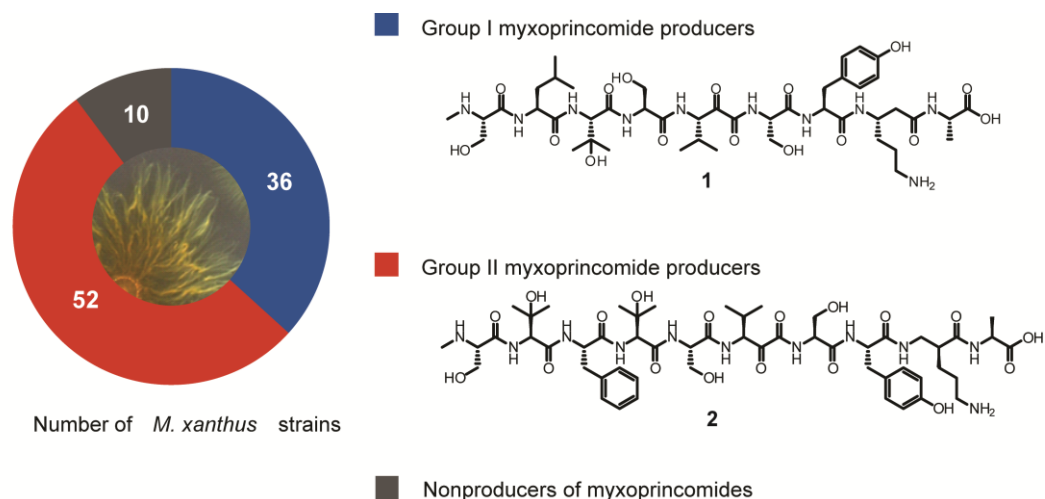
The supporting information is available online at:

[http://onlinelibrary.wiley.com/store/10.1002/anie.201106305/asset/supinfo/anie\\_201106305\\_sm\\_miscellaneous\\_information.pdf?v=1&s=d6c3b5184ef7ed27dd47db6aecbd5d0632e0876b](http://onlinelibrary.wiley.com/store/10.1002/anie.201106305/asset/supinfo/anie_201106305_sm_miscellaneous_information.pdf?v=1&s=d6c3b5184ef7ed27dd47db6aecbd5d0632e0876b)

## 4 The molecular diversity of the myxoprincomides: linear peptides assembled by an aberrant NRPS/ PKS hybrid synthetase

Myxoprincomides are linear peptides produced by *Myxococcus xanthus* constructed by the NRPS/PKS megaenzyme Mxp. Mxp is the largest, single, continuous hybrid synthetase reported from myxobacteria known to date which is capable of assembling an entire structural backbone.

The first reported myxoprincomide structure is that of myxoprincomide-c506 (**1**) from *M. xanthus* DK1622. Independent investigation of the secondary metabolome of another *M. xanthus* strain (DK897) uncovered a group of myxoprincomide analogs represented by myxoprincomide-c580 (**2**). Based on the mass spectrometric data obtained from the pilot study on the hidden secondary metabolome of 98 *M. xanthus* isolates, the production of these two groups of structural analogues is determined to be mutually exclusive.<sup>[1]</sup> (Fig 4.1) Consequently, it can be surmised that the corresponding hybrid synthetases Mxp<sub>1622</sub> and its heterofunctional homolog Mxp<sub>897</sub> are equally exclusory within the species. Group I is designated to the myxoprincomides produced by Mxp<sub>1622</sub> which is represented by myxoprincomide-c506 and the myxoprincomides assembled by the Mxp<sub>897</sub> pathway (myxoprincomide-c580 and structural variants) belong to the group II myxoprincomides. This chapter delves into the molecular diversity of the group I myxoprincomides. Unlike myxoprincomide-c506 and the group II myxoprincomides, for which the masses were uncovered during the *M. xanthus* secondary metabolome pilot study, the group I myxoprincomide derivatives to be presented in this chapter evaded detection by LC-MS mainly due to their low titers under standard conditions for *M. xanthus* cultivation. These structural variants were identified during the course of optimizing the isolation process for myxoprincomide-c506, an *M. xanthus* compound already considered to be produced in low abundance.



**Figure 4.1:** Tally of group I and group II myxoprincomide producers. The data is based on the pilot study on the secondary metabolomes of 98 *M. xanthus* isolates.<sup>[1]</sup> Group I myxoprincomides are represented by myxoprincomide-c506 (1) and group II myxoprincomides are represented by myxoprincomide-c580 (2). The image of *M. xanthus* was adapted from Velicer *et al.*, 2003.<sup>[2]</sup>

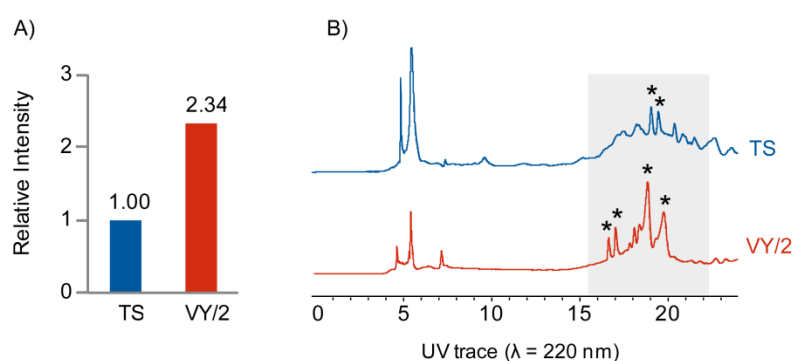
### 4.1 Optimization of myxoprincomide production and identification of 15 additional structural variants

#### 4.1.1 Optimization of myxoprincomide production

Myxoprincomide-c506 was purified from a cultivation of the optimized strain *M. xanthus* A2.c506 in 300 L PS medium (Section 2.19.4.1). PS medium is highly similar to TS medium, the medium which was previously identified as the optimal culture media for myxoprincomide production. PS medium differs from TS medium in that it uses peptone instead of tryptone as nitrogen source. Nevertheless, the chromatographic profile of the extract grown in either media are highly similar (data not shown).

Although the isolation of myxoprincomide-c506 from the PS medium cultivation was successful, the process was both labor and resource intensive. The main culprit is the overwhelming amount of polar media compounds which adsorb to the XAD-16 resin and coelute with myxoprincomide-c506. To illustrate, from ~25 g of lyophilized methanol crude extract derived from 100 L of culture broth, the product yield was ~1 mg after five chromatographic steps. Whereas the amount of purified compound was sufficient for structural elucidation by NMR, the time and resources spent in separating myxoprincomide-c506 from the media compounds deemed the process inefficient. Therefore an alternative cultivation condition was required to reduce the amount of interfering media compounds.

A viable alternative was the growth of the optimized strain A2.c506 in VY/2 medium. VY/2 medium is a complex medium long used for myxobacteria cultivation.<sup>[3]</sup> Chromatographic profiles of crude extracts from the cultivation of *Chondromyces* strains in VY/2 medium revealed that the components of this medium were observed minimally during sample work up. (R. Garcia & S. Hüttel, personal communication) Semi-quantitative analysis of myxoprincomide-c506 production in VY/2 and TS media revealed that myxoprincomide-c506 is produced two-fold more in VY/2 than in TS medium, with the added benefit of a reduced matrix effect on the chromatographic profile. (Fig. 4.2)



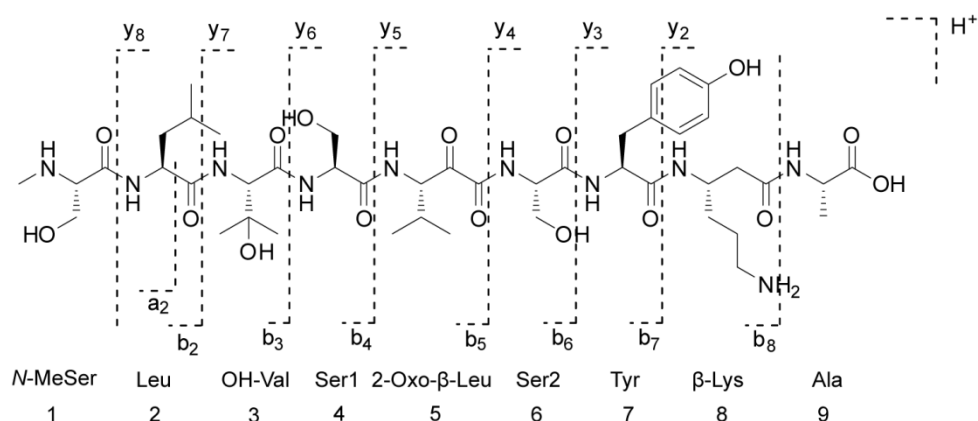
**Figure 4.2:** A) Comparison of the myxoprincomide-c506 production in TS and VY/2 media. B) UV traces of the crude extracts obtained from the cultivation of *M. xanthus* A2.c506 in TS and VY/2 media as analyzed using the HPLC method detailed in Section 2.20.1. Marked in asterisks (\*) are the peaks corresponding to myxoprincomide-c506.

Using VY/2 as cultivation medium, a 9-L scale fed batch fermentation yielded 0.5 mg of purified myxoprincomide-c506 with a simplified isolation process. (Section 2.20.2)

#### 4.1.2 Identification of myxoprincomide variants

The improved myxoprincomide-c506 production and sample work up facilitated the acquisition of MS<sup>2</sup> data which unambiguously identified 15 additional myxoprincomide variants. (Table 4.1)

The relationship to myxoprincomide-c506 was determined by two peptide sequence tags (PST) obtained from tandem MS analysis. PST A consists of the product ions 187 *m/z*, 215 *m/z*, 330 *m/z*, and 417 *m/z* which are consistent with the *a*<sub>2</sub>, *b*<sub>2</sub>, *b*<sub>3</sub>, and *b*<sub>4</sub> fragments from the fragmentation of myxoprincomide-c506. PST B includes 797 *m/z*, 682 *m/z*, and 595 *m/z* which correspond to *y*<sub>7</sub>, *y*<sub>6</sub>, and *y*<sub>5</sub> from myxoprincomide-c506. (Fig. 4.3 & S4.2)



**Figure 4.3:** Myxoprincomide-c506 with the y- and b- fragment ions marked. The shorthand form and position number for each amino acid residue are listed below the structure and will be used accordingly in the discussion.

The smallest and one of the largest myxoprincomide variants, named myxoprincomide-c534 and myxoprincomide-c708 respectively, were isolated and structurally elucidated by NMR and mass spectrometry analysis. Structures are proposed for nine structural variants based on *de novo* peptide sequencing using tandem MS data and feeding of labeled substrates. Four variants (c528, c535, c555, and c584) were identified based on their peptide sequence tags. However, due to their low titers at 50-ml cultivation scale, the volume in which the feeding of isotope-labeled amino acid substrates was carried out, structures could not be proposed for these derivatives.



**Table 4.1:** List of Group I myxoprincomides variants. The variants are further subdivided to Group I-A (short myxoprincomides), Group I-B (large myxoprincomides with PST A), Group I-C (large myxoprincomides with PST B), Group I-D (large myxoprincomides with PST A but lacking the 417 Da fragment ion). \* has identical  $y_2$ - $y_{10}$  product ions as myxoprincomide-c708.

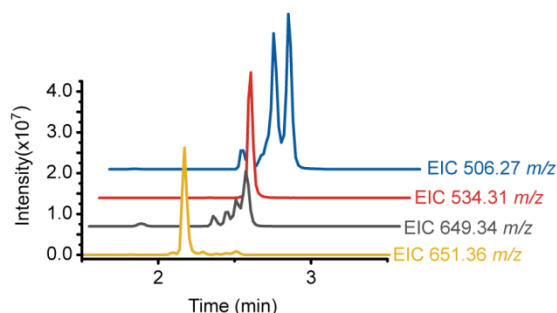
Group I-A						
Myxo ID	Cmpd	$m/z$ [M+H] <sup>+</sup> (measured)	Predicted Molecular Formula	$m/z$ (calc'd)	$\Delta$ ppm	PST
2668	c534	534.3134	C <sub>23</sub> H <sub>44</sub> O <sub>9</sub> N <sub>5</sub>	534.3133	0.085	A
2669	c649	649.3395	C <sub>27</sub> H <sub>49</sub> O <sub>12</sub> N <sub>6</sub>	649.3402	-1.228	A
2670	c651	651.3556	C <sub>27</sub> H <sub>51</sub> O <sub>12</sub> N <sub>6</sub>	651.3559	-1.746	A
Group I-B						
Myxo ID	Cmpd	$m/z$ [M+2H] <sup>2+</sup> (measured)	Predicted Molecular Formula	$m/z$ [M+2H] <sup>2+</sup> (calc'd)	$\Delta$ ppm	PST
2309	c506	506.2712 <sup>2+</sup>	C <sub>45</sub> H <sub>76</sub> N <sub>10</sub> O <sub>16</sub>	506.2721	-0.130	A
2754	c513	513.2789 (2 <sup>+</sup> )	C <sub>46</sub> H <sub>78</sub> O <sub>16</sub> N <sub>10</sub>	513.2793 (2 <sup>+</sup> )	-0.905	A
2755	c528	528.2841 (2 <sup>+</sup> )	C <sub>47</sub> H <sub>80</sub> O <sub>17</sub> N <sub>10</sub>	528.2846	-1.490	A
2767	c535	535.2915 (2 <sup>+</sup> )	C <sub>46</sub> H <sub>80</sub> O <sub>16</sub> N <sub>13</sub>	535.2917	0.671	A
2769	c591	591.8342 (2 <sup>+</sup> )	C <sub>44</sub> H <sub>74</sub> O <sub>17</sub> N <sub>10</sub> S	591.8337	0.655	A
2738	c562	562.8124 (2 <sup>+</sup> )	C <sub>51</sub> H <sub>87</sub> O <sub>17</sub> N <sub>11</sub>	562.8135	-0.465	A
2671	c708	708.3928 (2 <sup>+</sup> )	C <sub>66</sub> H <sub>108</sub> O <sub>20</sub> N <sub>14</sub>	708.3947	-0.363	A
Group I-C						
Myxo ID	Cmpd	$m/z$ [M+2H] <sup>2+</sup> (measured)	Predicted Molecular Formula	$m/z$ [M+2H] <sup>2+</sup> (calc'd)	$\Delta$ ppm	PST
2739	c515	515.2494 (2 <sup>+</sup> )	C <sub>44</sub> H <sub>74</sub> O <sub>16</sub> N <sub>10</sub> S	515.2496	-1.222	B
2762	c523a	523.2466(2 <sup>+</sup> )	C <sub>44</sub> H <sub>74</sub> O <sub>17</sub> N <sub>10</sub> S	523.2468	0.226	B
2740	c523b	523.2632 (2 <sup>+</sup> )	C <sub>48</sub> H <sub>74</sub> O <sub>16</sub> N <sub>10</sub>	523.2637	-0.792	B
2842	c717	717.3715 (2 <sup>+</sup> )	C <sub>65</sub> H <sub>106</sub> O <sub>20</sub> N <sub>14</sub> S	717.3708	0.812	*
Group I-D						
Myxo ID	Cmpd	$m/z$ [M+2H] <sup>2+</sup> (measured)	Predicted Molecular Formula	$m/z$ [M+2H] <sup>2+</sup> (calc'd)	$\Delta$ ppm	PST
2770	c555	555.8055 (2 <sup>+</sup> )	C <sub>47</sub> H <sub>80</sub> O <sub>17</sub> N <sub>10</sub>	555.8050	1.484	A - 417
2768	c584	584.8265 (2 <sup>+</sup> )	C <sub>44</sub> H <sub>74</sub> O <sub>16</sub> N <sub>10</sub> S	584.8259	0.945	A - 417

#### 4.1.2.1 Group I-A: identification and characterization of myxoprincomide-c534, -c649, and -c651

LC-MS/MS<sup>2</sup> analysis of the size exclusion chromatography (SEC) fraction eluting after myxoprincomide-c506 revealed 3 compounds with monoisotopic masses  $m/z$  [M+H]<sup>+</sup> = 534.3 Da, 649.3 Da, and 651.4 Da whose MS<sup>2</sup> fragmentation pattern include PST A.

## 4 Myxoprincomides

Reverse phase HPLC-MS analysis revealed that the retention times of these compounds fall within the time range when myxoprincomide-c506 elutes. (Fig. 4.4)



**Figure 4.4:** Chromatographic profiles of myxoprincomide-c506, c534, c649, and c651 as produced in VY/2 medium

Compound c534 elutes as a single peak with a monoisotopic mass of  $m/z$   $[M+H]^+ = 534.3134$  corresponding to a predicted molecular formula of  $C_{23}H_{44}N_5O_9$  ( $\Delta ppm = 0.085$ ) as analyzed by high resolution ESI-MS.

MS<sup>2</sup> fragmentation of c534 revealed  $y$  ions whose neutral losses correspond to the five N-terminal amino acid residues in myxoprincomide-c506. (Fig. S4.3) Feeding with  $^{13}C_3,^{15}N$ -L-serine,  $d_3$ -L-leucine, and  $d_8$ -L-valine confirmed the incorporation of two serine, one leucine, and 2 valine residues. (Fig. S4.3) An observed loss of a deuterium atom from one of the incorporated  $d_8$ -L-valine residues could correspond to the displacement of the deuterium atom at  $C\beta$  by a  $-OH$  group to form L- $\beta$ -OH-valine. (Fig. S4.1) Growth in the presence of L-methionine-(methyl- $^{13}C$ ) showed the incorporation of the labeled methyl group carbon attributed to the N-methylation of L-serine, possibly derived from an intracellular pool of (methyl- $^{13}C$ )-S-adenosyl- $^{13}C$ -methionine.

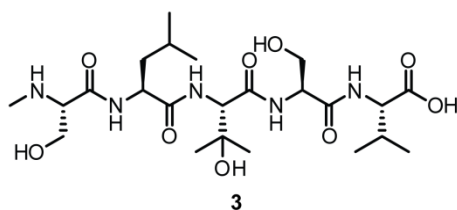
To confirm the structure, 1.1 mg of c534 was purified and analyzed by one- and two-dimensional (1D & 2D) NMR experiments ( $^1H$ ,  $^{13}C$ ,  $[^1H,^1H]$ -HSQC, and  $[^1H-^{13}C]$ -HMBC) in DMSO- $d_6$ . (Table 4.2 and Fig. S4.14-18)

**Table 4.2:** NMR spectroscopic data for myxoprincomide-c534 (DMSO- $d_6$ ). <sup>a</sup> Recorded at 175 MHz; referenced to residual DMSO- $d_6$  at  $\delta$  39.51 ppm. <sup>b</sup> Recorded at 700 MHz; referenced to residual DMSO- $d_6$  at  $\delta$  2.50 ppm. <sup>c</sup> Proton showing HMBC correlation to indicated carbon

	$\delta_C^a$	$\delta_H^b$ (J in Hz)	HMBC <sup>c</sup>
N-MeSer			
1	166.0		
2	61.9	3.83 m	1, 3
3	59.6	3.83, 3.68	2
NH <sub>2</sub> <sup>+</sup>		8.43 8.04 t	2
Nme	31.3	2.51	
OH		5.48 br s	
Leu			
1	171.2		
2	51.2	4.48 q	1, 3, 4, 1 <sub>NmeSer</sub>
3	40.1	1.49 t	1, 2, 4, 5, 6
4	23.9	1.62 m	2, 3, 5, 6
5	21.1	0.87 d (6.1)	3, 4, 6
6	22.9	0.88 d (6.7)	3, 4, 5
NH		8.68 d (7.7)	1 <sub>NmeSer</sub>
OH-Val			
1	169.4		
2	59.5	4.43 d (9.2)	1, 3, 4, 5, 1 <sub>Leu</sub>
3	71.7		
4	27.6	1.12 s	2, 3, 4
5	24.4	1.09 s	2, 3, 5
NH		7.95 d (8.2)	1 <sub>Leu</sub>
OH		5.01 s	2, 3, 4, 5
Ser			
1	170.0		
2	54.8	4.36 m	1, 3, 1 <sub>OH-Val</sub>
3	59.4	3.68, 3.59 m	1, 2
NH		7.95 d (8.2)	1 <sub>OH-Val</sub>
OH		4.96 br	
Val			
1	172.5		
2	57.2	4.12 dd (6.2, 6.1)	1, 3, 4, 5, 1 <sub>Ser1</sub>
3	29.6	2.03 m	1, 2, 4, 5
4	18.9	0.87 d (6.5)	2, 3, 4
5	17.9	0.86 d (6.5)	2, 3, 5
NH		7.90 d (8.3)	1 <sub>Ser</sub>

Using the NMR spectra of myxoprincomide-c506 as reference, the observed <sup>1</sup>H and <sup>13</sup>C chemical shifts from the c534 NMR spectra were compared and subsequently shown to be in good agreement with those of amino acids 1-5 of myxoprincomide-c506. (Fig. 4.5) The only clear difference observed is C1 from L-valine which has a <sup>13</sup>C chemical shift of  $\delta$

172.5 ppm in c534 but is  $\delta$  196.4 ppm in myxoprincomide-c506. The deviation is expected as the L-valine C1 from c534 is part of the -COOH terminus while the same carbon in myxoprincomide-c506 belongs is C2 of the L-valine-derived 2-oxo-leucine moiety. The absolute configuration of the amino acid residues were determined by advanced Marfey's method and myxoprincomide-c534 was determined to contain only L- amino acids as previously observed in myxoprincomide-c506.

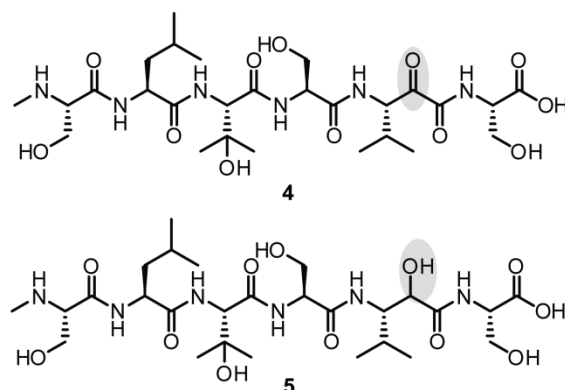


**Figure 4.5:** Structure of myxoprincomide-c534 (3)

Compound c649 and c651 were analyzed by HR ESI-MS and were determined to have a monoisotopic mass of  $m/z$   $[M+H]^+ = 649.3396$  Da (predicted formula of  $C_{27}H_{49}N_6O_{12}$  ( $\Delta$ ppm = -1.228)) and  $m/z$   $[M+H]^+ = 651.3556$  Da (predicted formula  $C_{27}H_{51}N_6O_{12}$  ( $\Delta$ ppm = -1.746)) respectively. Compound c649 elutes as multiple peaks akin to myxoprincomide-c506 while compound c651 elutes slightly earlier as a single peak. (Fig 4.4)

Tandem MS analysis of both c649 and c651 clearly generated PST A implying that the first four amino acids are comprised of N-MeSer, Leu, OH-Val, and Ser, and that the difference between these two compounds lie on the two C-terminal amino acid residues. Feeding with  $d_8$ -L-valine and  $^{13}C_3,^{15}N$ -L-serine showed that both compounds contain two L-valine derived residues and three L-serine residues. Neutral losses to form the  $b_5$  ion 544.22  $m/z$  from c649 and  $b_5$  ion 546.27  $m/z$  from c651 indicate a loss of an L-serine residue at the C-terminus (Fig. S4.4 & S4.5)

The subsequent neutral loss to form the  $b_4$  ion 417.15  $m/z$  indicated that the difference between c649 and c651 lies between Ser2 and Ser3 of the molecules. Using the structural information from the group II myxoprincomides from *M. xanthus* DK897 where 2-oxo and 2-hydroxy structural features were observed (O. Revemann, Dissertation, Universität des Saarlandes, 2012), we propose that c649 is the shunt product with the 2-oxo moiety is retained and c651 contains a 2-hydroxy group in lieu of the 2-oxo functional group. (Fig. 4.6)

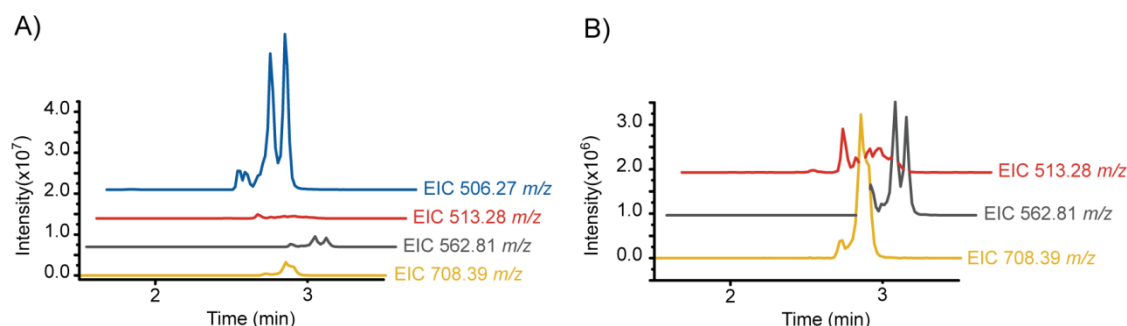


**Figure 4.6:** Proposed structure for myxoprincomide-c649 (4) and for myxoprincomide-c651 (5). The proposed stereochemistry is based on the absolute configuration of the representative structure myxoprincomide-c506.

Myxoprincomide-c534, -c649, and c-651 were confirmed to be shunt products released from the Mxp<sub>1622</sub> assembly line. The possibility that these 3 shorter myxoprincomides are post assembly degradation products was negated when all 3 compounds were detected in the crude extract of DK1622 mutant strains which could not produce the necessary  $\beta$ -lysine precursor necessary for the biosynthesis of the larger myxoprincomides (e.g. myxoprincomide-c506). (Section 4.2.1)

#### 4.1.2.2 Group I-B : Identification and characterization of c513, c562 and c708

Compounds c513, c562 and c708 were determined to be related to myxoprincomide-c506 after LC-MS/MS<sup>2</sup> analysis of the SEC fraction eluting before myxoprincomide-c506. The MS<sup>2</sup> spectra of c513, c562, and c708 all reflect PST A, establishing the relationship to myxoprincomide-c506.

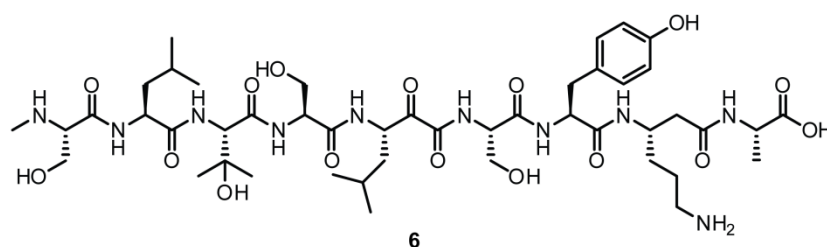


**Figure 4.7:** Chromatographic profiles of myxoprincomide-c506, -c513, -c562, and -c708 as produced in VY/2 medium (A). Magnified EIC of c513, c562, and c708

Compound c513 has a monoisotopic mass of  $m/z$   $[M+2H]^{2+} = 513.2789$  Da with a predicted formula of  $C_{46}H_{78}O_{16}N_{10}$  ( $\Delta ppm = -0.905$ ) which corresponds to an additional

methylene group (-CH<sub>2</sub>-) in comparison to myxoprincomide-c506. Analysis of the d<sub>8</sub>-L-valine labeled c513 revealed the incorporation of a labeled L-β-OH-valine but not a labeled L-valine. (Fig. S4.6) This implies that only OH-Val is incorporated into c513 and that the second valine is replaced by another amino acid which we believe to be a leucine as an additional L-leucine is observed in the d<sub>3</sub>-L-leucine-labeled c513.

MS<sup>2</sup> analysis of c513 disclosed that the neutral losses between the product ions are congruent with most of the monomers observed in myxoprincomide-c506 with the exception of  $y_5 \rightarrow y_4$ . The neutral loss between  $y_5$  and  $y_4$  from c513 contains an additional 14 Da in comparison to the same position in myxoprincomide-c506 consistent with the Val→Leu hypothesis. The proposed structure is presented in Figure 4.8.

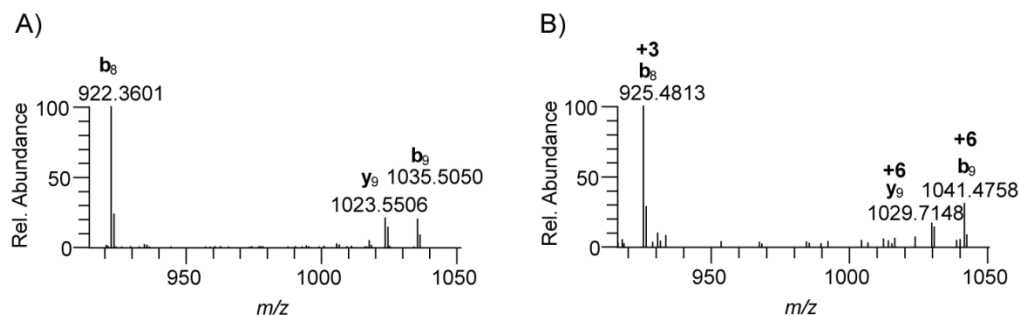


**Figure 4.8:** Proposed structure for myxoprincomide-c513 (6). The proposed stereochemistry is based on the absolute configuration of the representative structure myxoprincomide-c506.

Based on HR ESI-MS, compound c562 has a monoisotopic mass of  $m/z$   $[M+H]^+ = 562.8135^{2+}$  and a predicted formula of  $C_{51}H_{87}N_{11}O_{17}$  ( $\Delta ppm = -0.465$ ). The mass difference between c562 and myxoprincomide-c506 is 113.08 Da implying an incorporation of an additional L-leucine. The incorporation of two  $d_3$ -L-leucine residues in c562 confirms this hypothesis. (Fig. S4.7)

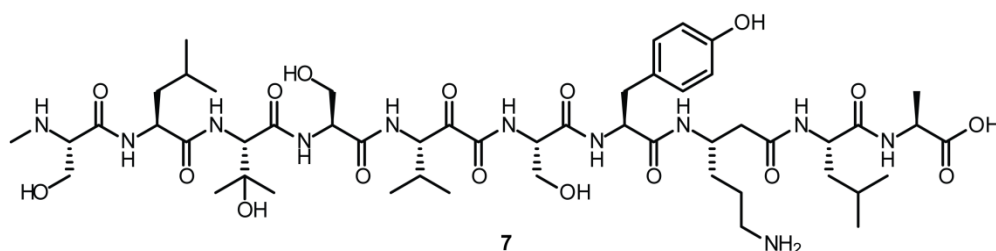
Examination of the b ions from the isotope-labeled c562 enabled the determination of the position of the second L-leucine in the molecule. (Fig. 4.9) The b<sub>9</sub> ion 1035.61 *m/z* indicate the loss of L-alanine. This b<sub>9</sub> ion is observed to contain two *d*<sub>3</sub>-L-leucine residues.

The  $b_8$  ion reflected a mass difference from  $b_9$  consistent with a L-leucine residue. The  $b_8$  ion from the  $d_3$ -L-leucine labeled c562 appears to contain only one  $d_3$ -L-leucine thereby confirming that L-leucine was indeed the amino acid lost from  $b_9$  to  $b_8$ . The subsequent  $b_7$ ,  $b_6$ , and  $b_5$  ions were consistent with the loss of a L- $\beta$ -lysine, L-tyrosine, and L-serine respectively.



**Figure 4.9:** Observed  $b_8$ ,  $y_9$  and  $b_9$  product ions of the unlabeled (A) and  $d_3$ -L-leucine-labeled myxoprincomide-c562.

Given the tandem MS data from unlabeled and isotope-labeled c562, a structure for myxoprincomide-c562 is proposed. (Fig. 4.10)



**Figure 4.10:** Proposed structure for myxoprincomide-c562 (7). The proposed stereochemistry is based on the absolute configuration of the representative structure myxoprincomide-c506.

Determination of c708 as a myxoprincomide variant via the identification of PST A initially led us to believe that we obtained the intact myxoprincomide with a hypothetical acyl starter unit preceding the *N*-MeSer residue. However,  $^{13}\text{C}$ -acetate labeled c708 only indicated the incorporation of one C2 carbon from acetate, similar to myxoprincomide-c506, eliminating the possibility of an acetate-derived acyl starter chain. Analysis of the c708 containing isotope labeled amino acids revealed that c708 contains extra leucine and phenylalanine-derived residues in comparison to myxoprincomide-c506. Assuming that L-tyrosine is the additional phenylalanine derived residue, the only unaccounted mass difference is 128 Da which could correspond to a lysine or a  $\beta$ -lysine. Fragment ions  $a_2$  and  $b_2 - b_8$  from c708 and myxoprincomide-c506 are identical, denoting that this portion of the two compounds is putatively identical. (Fig. S4.8) Subsequent neutral losses from  $b_9$  to  $b_{11}$  appear to

## 4 Myxoprincomides

correspond to tyrosine, lysine/  $\beta$ -lysine, and leucine. The difference between  $M+H^+$  and  $b_{11}$  is equivalent to a loss of alanine.

1D and 2D NMR experiments were carried out on 0.5 mg of pure c708. (Fig. S4.19-24) The NMR spectra were recorded in the same solvent and instrument as that used for the structural elucidation of myxoprincomide-c506 to facilitate comparison. Examination of the -NH region of the  $^1H$  spectra indicated 3 NH groups ( $\delta$  8.16 ppm,  $\delta$  7.59 ppm,  $\delta$  7.98 ppm) not identified in myxoprincomide-c506. The proton with a chemical shift of  $\delta$  8.16 ppm was determined to belong to a tyrosine residue as determined by [ $^1H$ ,  $^{13}C$ ]-HSQC and [ $^1H$ ,  $^{13}C$ ]-HMBC. Using TOCSY couplings of the -NH protons with chemical shifts  $\delta$  7.59 ppm,  $\delta$  7.98 ppm were determined to belong to  $\beta$ -lysine and leucine respectively. Assignment of the rest of the c708 signals was facilitated by their identity to those observed in the NMR spectra of myxoprincomide-c506. (Table 4.3)

**Table 4.3:** NMR spectroscopic data for myxoprincomide-c708 (DMSO- $d_6$ ). <sup>a</sup> Recorded at 175 MHz; referenced to residual DMSO- $d_6$  at  $\delta$  39.51 ppm. <sup>b</sup> Recorded at 700 MHz; referenced to residual DMSO- $d_6$  at  $\delta$  2.50 ppm. <sup>c</sup> Proton showing HMBC correlation to indicated carbon. <sup>d</sup> Proton showing ROESY correlation to indicated proton, \* not assigned

	$\delta_C^a$	$\delta_H^b$ (J in Hz)	HMBC <sup>c</sup>	ROESY <sup>d</sup>
NmeSer				
1	166.2			
2	61.8	3.86 br s	1, 3	Nme, NH <sub>Leu</sub>
3	59.4	3.83, 3.70m	2	Nme, NH <sub>2</sub> <sup>+</sup> , NH <sub>Leu</sub>
NH <sub>2</sub> <sup>+</sup>		8.84, 8.76 br s		2, Nme
Nme	31.1	2.52	2	
OH		5.49 br s		
Leu1				
1	171.6			
2	51.2	4.50 dt (8.8, 6.3)	1, 3, 1 <sub>NmeSer</sub>	3, 4, 5, NH, NH <sub>OH-Val</sub>
3	40.3	1.48 m	1, 2, 4, 5	2, 5, NH
4	24.1	1.63 m	2, 3, 5, 6	2, NH
5	21.1	0.87 d (6.6)	3, 6	
6	22.8	0.89 d (6.6)	3, 4, 5	
NH		8.70 d (8.3)	1 <sub>NmeSer</sub>	2, 3, 4
OH-Val				
1	169.2			
2	59.5	4.43 d (9.1)	1, 3, 4, 5, 1 <sub>Leu</sub>	4
3	71.9			
4	27.3	1.13 s	2, 3, 5	2, NH
5	24.5	1.08 s	2, 3, 4	
NH		7.99 d (8.2)	1 <sub>Leu1</sub>	4, 1 <sub>Leu1</sub>
OH		*		
Ser1				
1	170.8			

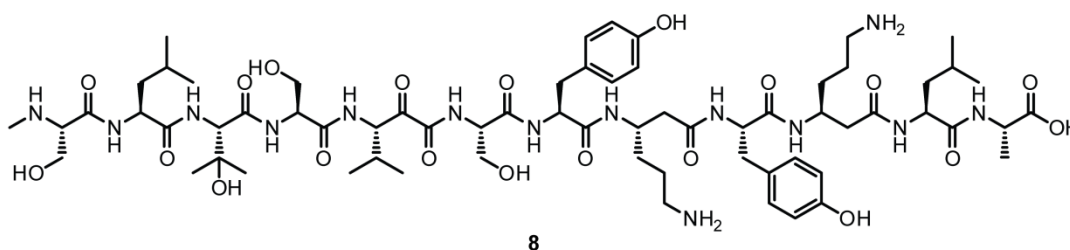


2	55.3	4.38 m	1, 3, 1 <sub>OH-Val</sub>	3, NH
3	61.6	3.37 3.58 m	1, 2	NH <sub>OH-Val</sub>
NH		8.00 d (8.1)	1 <sub>OH-Val</sub>	2
OH		5.20 br		
2-oxo- $\beta$ -Leu				
1	172.1			
2	196.9			
3	57.6	5.14 dd (5.4, 5.5)	2, 4, 5, 6, 1 <sub>Ser1</sub>	4, 5, 6, NH, NH <sub>Ser2</sub>
4	29.5	2.17 m	3, 4, 5	3, 5, 6
5	19.7	0.87 d (6.6)	3, 4, 6	3, 4
6	17.5	0.81 d (6.8)	3, 4, 5	3, 4, NH
NH		8.05 d (8.0)	1 <sub>Ser1</sub>	3, 6, 2 <sub>Ser1</sub>
Ser2				
1	168.8			
2	50.3	4.33 m	3, 1 <sub>2-oxo-<math>\beta</math>-Leu</sub>	3, NH, NH <sub>Tyr</sub>
3	61.6	3.62, 3.57 m	1, 2	1, NH
NH		8.33 d (7.9)	1 <sub>2-oxo-<math>\beta</math>-Leu</sub>	2, 3, 3 <sub>2-oxo-<math>\beta</math>-Leu</sub>
OH		*		
Tyr1				
1	170.5			
2	54.8	4.35 m	3	3, 2', 6' NH <sub><math>\beta</math>-Lys1</sub>
3	36.9	2.87 dd (14.0, 4.3)	1, 2, 1', 2'	2, 2', 6'
		2.68 dd (14.1, 9.0)	1, 2, 1', 2'	2', 6'
1'	127.9			
2', 6'	130.2	6.98 d (8.6)	3, 4'	2, 3
3', 5'	114.6	6.63 d (8.4)	4'	
4'	156.0			
NH		8.16 d (7.5)	1 <sub>Ser2</sub>	2, 3, 2', 3', 2 <sub>Ser2</sub>
OH		9.16 s		
$\beta$ -Lys1				
1	169.5			
2	40.1	2.12 br d (6.2)	1, 3	3, NH, NH <sub>Tyr2</sub>
3	45.2	3.92 m		2, 4, 5, 6, NH
4	29.9	1.35, 1.09 m	2, 3, 6	3, NH <sub>2</sub> , NH
5	23.4	1.35, 1.45 m	3, 6	6, NH <sub>2</sub>
6	38.3	2.63 br	4, 5	5, NH <sub>2</sub>
NH <sub>2</sub>		7.59 m	6	5, 6
NH		7.80 d (8.8)	1 <sub>Tyr1</sub>	2, 3, 4, 2 <sub>Tyr1</sub>
Tyr2				
1	170.8			
2	54.3	4.38 m	3	3, 2', 6' NH <sub><math>\beta</math>-Lys2</sub>
3	37.0	2.82 dd (13.5, 3.9)	1, 2, 1', 2'	2, 2', 6'
		2.59	1, 2, 1', 2'	2', 6'
1'	128.1			
2', 6'	130.0	6.99 d (8.6)	3, 4'	2, 3
3', 5'	114.6	6.63 d (8.4)	4'	
4'	156.0			
NH		8.04 d (8.4)	1 <sub><math>\beta</math>-Lys1</sub>	2, 3, 2', 6', 2 <sub><math>\beta</math>-Lys1</sub>
OH		9.16 s		
$\beta$ -Lys2				
1	170.0			

## 4 Myxoprincomides

2	40.1	2.20 br d (6.2)	1, 3, 4	NH, NH <sub>Leu2</sub>
3	45.3	4.02 br		NH <sub>Leu2</sub>
4	30.3	1.46, 1.36 m	2, 3, 6	6, NH
5	23.6	1.46, 1.52 m	2, 6	6, NH <sub>2</sub>
6	38.4	2.73 dd (12.0, 7.0)	4, 5	5, NH <sub>2</sub>
NH <sub>2</sub>		7.64 m	5, 6	5, 6
NH		7.89 d (8.3)	1 <sub>Tyr2</sub>	2, 2 <sub>Leu2</sub>
Leu2				
1	172.1			
2	50.3	4.33 m	1	NH <sub>Ala</sub>
3	40.8	1.43 m	1	2, 5, NH, NH <sub>Ala</sub>
4	23.8	1.62 m	2, 3, 5, 6	2, NH
5	21.8	0.84 d (6.6)	3, 4, 6	2
6	23.5	0.87 d (6.6)	3, 5	
NH		7.98 d (8.2)	1 <sub>β-Lys2</sub>	2 <sub>β-Lys2</sub>
Ala				
1	174.3			
2	47.1	4.13 t (7.3)	1, 3, 1 <sub>Leu2</sub>	3, NH, 2 <sub>Leu2</sub>
3	16.8	1.22 d (7.3)	1, 2	2, NH
NH		8.23 d (7.3)	2, 3, 1 <sub>Leu2</sub>	3, 2 <sub>Leu2</sub>

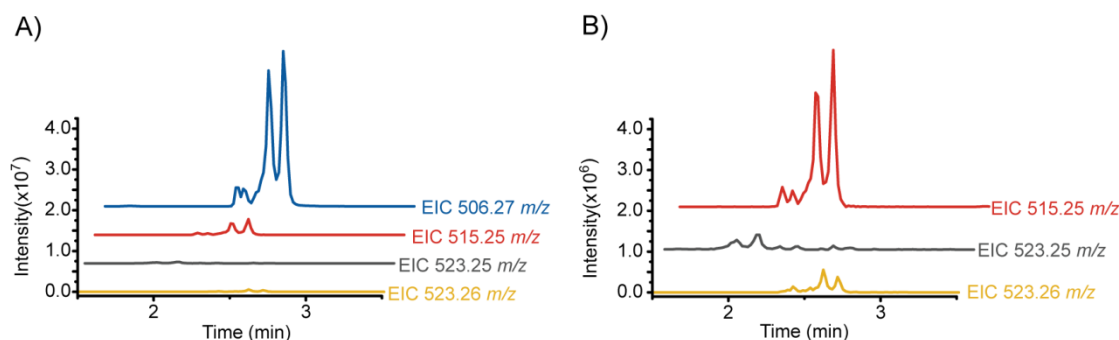
Using the [<sup>1</sup>H, <sup>13</sup>C]-HMBC connectivities, the correct sequence of the amino acids were determined. Myxoprincomide-c708 is comprised of L-amino acid subunits as determined by advanced Marfey's method. The final structure is shown in Figure 4.11.



**Figure 4.11:** Final structure of myxoprincomide-c708 (8)

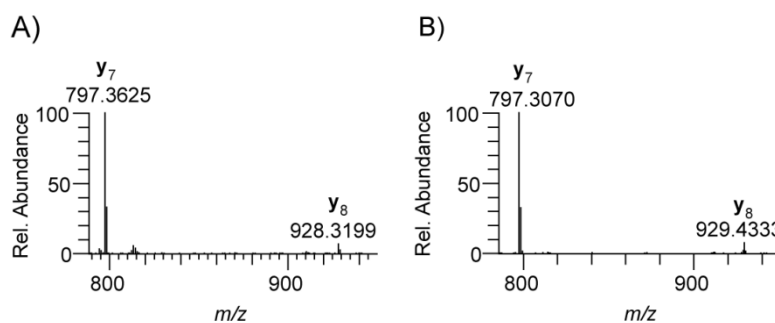
### 4.1.2.3 Group I-C: Identification and characterization of myxoprincomide-c515, -c523a, c523b

Group I-C myxoprincomides are characterized by the PST B. Common to all three compounds is the replacement of Leu at position 2, as confirmed by the lack of *d*<sub>3</sub>-L-leucine incorporation (Figure S4.9-11).



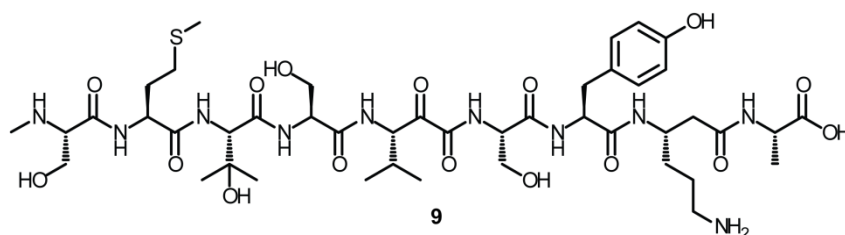
**Figure 4.12:** Chromatographic profiles of myxoprincomide-c506, -c515, -c523a, and -c523b as produced in VY/2 medium (A). Magnified EICs of c515, c523a, and c523b

The incorporation of two  $^{13}\text{C}$ -labeled methyl groups derived from L-methionine-( $^{13}\text{C}$ -methyl) could be observed in compound c515. (Fig. S4.10) One of these  $^{13}\text{C}$ -labeled methyl groups is correlated to the *N*-MeSer residue and is underpinned by the loss of one  $^{13}\text{C}$ -label after the neutral loss (101 Da) to form the  $y_8$  ion. The subsequent  $y_7$  product ion no longer holds a  $^{13}\text{C}$ -methyl label indicating the position of the second  $^{13}\text{C}$ -labeled methyl containing residue. (Fig. 4.13)

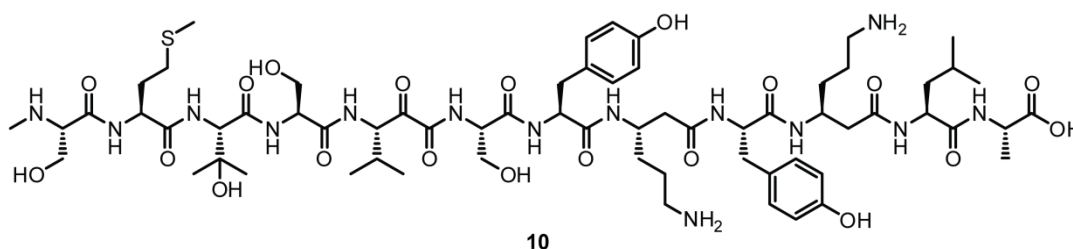


**Figure 4.13:** Observed  $y_7$  and  $y_8$  product ions from the fragmentation of unlabeled (A) and L-methionine-( $^{13}\text{C}$ -methyl)-labeled myxoprincomide-c515

The fragment ion  $y_8$  (928.32  $m/z$ ) is 131.04 Da larger than the  $y_7$  ion (747.36  $m/z$ ) which corresponds to the residue mass for methionine. The incorporation of an intact methionine accounts for the second  $^{13}\text{C}$ -labeled methyl group. The replacement of Leu by Met, a less hydrophobic amino acid, could explain why c515 elutes slightly earlier than myxoprincomide-c506 under standard LC-MS analysis (Fig. 4.14). A methionine-containing analog of myxoprincomide-c708 was also identified with a monoisotopic mass  $m/z$   $[\text{M}+2\text{H}]^{2+} = 717.3715$  Da (predicted formula:  $\text{C}_{65}\text{H}_{106}\text{O}_{20}\text{N}_{14}\text{S}$ ,  $\Delta\text{ppm} = 0.812$ ). (Fig. 4.15 & S4.12)

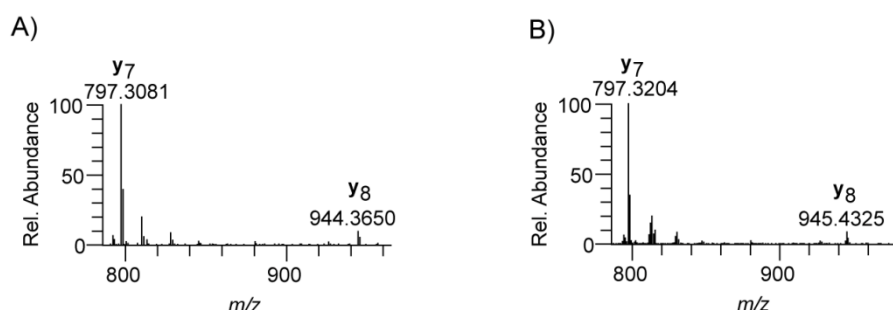


**Figure 4.14:** Proposed structure for myxoprincomide-c515 (9). The proposed stereochemistry is based on the absolute configuration of the representative structure myxoprincomide-c506.



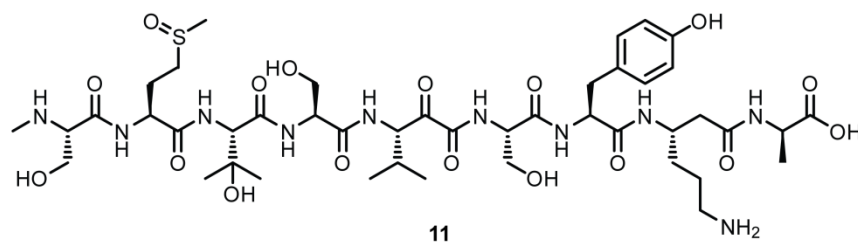
**Figure 4.15:** Proposed structure for myxoprincomide-c717 (10). The proposed stereochemistry is based on the absolute configuration of myxoprincomide-c708.

Similar to myxoprincomide-c515, analysis of the  $^{13}\text{C}$ -methyl labeled c523a revealed the incorporation of two  $^{13}\text{C}$ -methyl groups and their corresponding neutral losses indicate once again that both labeled carbon atoms are in the first two amino acid residues. (Fig. 4.16 and Fig. S4.10)



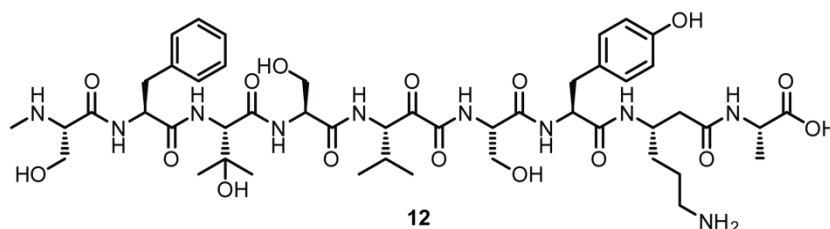
**Figure 4.16:** Observed  $y_7$  and  $y_8$  product ions from the fragmentation of unlabeled (A) and L-methionine-( $^{13}\text{C}$ -methyl)-labeled myxoprincomide-c523a

The difference between  $y_8$  and  $y_7$  is 147.04 Da which corresponds to the residue mass of methionine sulfoxide. The calculated monoisotopic mass ( $m/z$   $[\text{M}+2\text{H}]^{2+} = 523.2468$ ) for the proposed c523a structure (Fig. 4.17) is in very good agreement with the measured  $m/z$  value ( $m/z$   $[\text{M}+2\text{H}]^{2+} = 523.2466$ )



**Figure 4.17:** Proposed structure for myxoprincomide-c523a (11). The proposed stereochemistry is based on the absolute configuration of the representative structure myxoprincomide-c506.

Compound c523b, which has a 0.04 Da higher mass than c523a ( $m/z$   $[M+2H]^{2+} = 523.2633$  Da, predicted formula  $C_{48}H_{74}O_{16}N_{10}$ ,  $\Delta ppm = -0.792$ ), likely contains a phenylalanine in position 2. (Fig. 4.18) We propose this based on the detection of a doubly L-phenylalanine- $^{13}C_6$ -ring labeled c523b and the observed  $y_8 \rightarrow y_7$  loss of a phenylalanine residue (147.07 Da) from tandem MS analysis. (Fig. S4.11)



**Figure 4.18:** Proposed structure for myxoprincomide-c523b. The proposed stereochemistry is based on the absolute configuration of the representative structure myxoprincomide-c506.

## 4.2 Biosynthesis of the non-proteinogenic amino acid precursors L- $\beta$ -lysine and L- $\beta$ -hydroxyvaline

Observed from the group I myxoprincomides are two non-proteinogenic amino acids: L- $\beta$ -lysine, present in the larger group I myxoprincomides (MW > 1000 Da) and L- $\beta$ -OH-valine, observed in all variants. As genes encoding the enzymatic functions for the biosynthesis of these two unusual amino acids were not observed in the vicinity of the *mvp*<sub>1622</sub> (Fig 3.4) the enzymes responsible for the formation of L- $\beta$ -lysine and L- $\beta$ -OH-valine were hypothesized to be encoded by distinct genes.

### 4.2.1 Formation of $\beta$ -lysine

Beta-lysine is a non-proteinogenic amino acid found in nature as an intermediate in the anaerobic catabolism of L-lysine.<sup>[4]</sup> In some *Streptomyces* species,  $\beta$ -lysine acts as a

precursor to produce bioactive compounds such as tuberactinomycins (e.g. viomycin) streptothricins, and nourseothricin.<sup>[5,6]</sup> The transfer of the amino group from the C $\alpha$  to C $\beta$  is in these cases catalyzed by SAM-dependent 2,3-lysine aminomutases (2,3-LAM).<sup>[4]</sup>

The most well characterized 2,3-LAM is KamA which is involved in lysine metabolism in *Clostridium subterminale* SB4. X-ray crystallographic studies on KamA, where the substrate and cofactors are bound to the active center, have shown that catalysis requires a CxxxCxxC binding motif which encloses a [4Fe-4S] cluster, Lys<sub>337</sub> for the binding of pyridoxal pyrophosphate, and Arg<sub>134</sub>-Asp<sub>293</sub>-Asp<sub>330</sub> which interact with the carboxyl and amino groups of the lysine substrate.<sup>[7]</sup>

Annotation of the genes upstream and downstream of *mxp*<sub>1622</sub> did not reveal any aminomutase-encoding genes that could possibly catalyze the conversion of lysine to  $\beta$ -lysine. However, it is plausible that the 2,3-LAM encoding gene could be located somewhere else in the genome similar to the dispersed genetic loci of the aurachin biosynthetic pathway.<sup>[8]</sup>

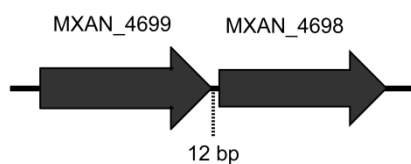
Based on their sequence similarity of their gene products to the amino acid sequence of KamA, VioP (1,335 bp, 455 aa), a 2,3-LAM involved in the biosynthesis of viomycin, and SttO (CAB67710, 1,201 bp, 403 aa) of the streptothricin biosynthetic pathway,<sup>[5]</sup> 3 genes, namely MXAN\_4699, MXAN\_4698, and MXAN\_6263 were identified as possible 2,3-LAM-encoding candidates. (Table 4.4 and Fig. S4.13) The gene products of these 3 *M. xanthus* genes contain the required residues for 2,3-LAM activity. (Fig. S4.13)

**Table 4.4:** Percent identity / similarity of putative *M. xanthus* DK1622 2,3-KAM encoding genes to VioP (Accession no. AAP92506.1)

Gene	Length (bp / aa)	Identity / Similarity (in %)
MXAN_4699	1,371 / 456	61 / 74
MXAN_4698	1,233 / 410	31 / 46
MXAN_6263	1,329 / 443	26 / 45

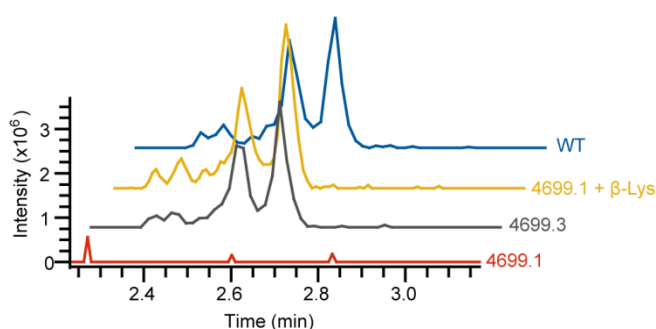
These three genes were individually inactivated and out of the three only the MXAN\_4699 knockout mutant strain 4699.1 was deficient in myxoprincomide-c506 production. Whereas a polar effect on the downstream MXAN\_4698 gene (Fig. 4.19) could in theory contribute to the loss of  $\beta$ -lysine production, a synergistic relationship between the two proteins for the biosynthesis of the myxoprincomide  $\beta$ -lysine precursor

is unlikely as independent inactivation of MXAN\_4698 did not impede myxoprincomide-c506 production.



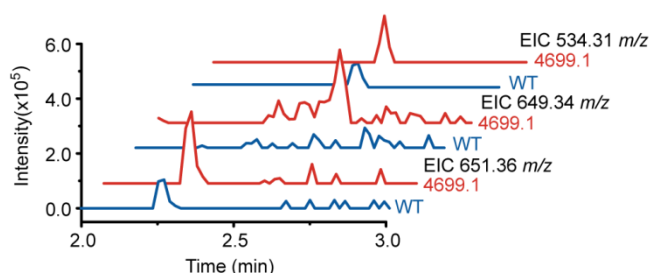
**Figure 4.19:** Gene organization of MXAN\_4699 and MXAN\_4698 showing the intergenic region with a length of 12 bp.

To further confirm the role of MXAN\_4699 as the 2,3-LAM involved in myxoprincomide biosynthesis, genetic complementation of MXAN\_4699 (mutant strain 4699.3) by insertion of the native gene into the *attB* phage attachment site of the knockout strain 4699.1 was constructed. Alternatively, the 4699.1 production culture was supplemented with 1 mM L- $\beta$ -lysine. Both approaches restored myxoprincomide-c506 production. (Fig. 4.20)



**Figure 4.20:** Extracted ion chromatograms for myxoprincomide-c506 from the wildtype, 2,3-KAM knockout mutant 4699.1, the complementation mutant 4699.3, and the L- $\beta$ -lysine-fed 4699.1 mutant strain.

Interestingly, the assembly of the group I-A myxoprincomides were unaffected by the inactivation of MXAN\_4699, thereby suggesting that these intermediates are released from Mxp<sub>1622</sub> and are not products of post-assembly degradation. (Fig. 4.21)



**Figure 4.21:** Extracted ion chromatograms of the group I-A myxoprincomides from the wildtype and 2,3-KAM knockout mutant 4699.1.

### 4.2.2 Formation of L- $\beta$ -hydroxyvaline

One of the prominent features of the myxoprincomides is the unusual L- $\beta$ -hydroxyvaline (OH-Val) situated consistently at position 3 of the group I myxoprincomides. In group II myxoprincomides, this non-proteinogenic amino acid is even found twice, one at position 2 and the other position 4. (O. Revermann, Dissertation, Universität des Saarlandes, 2012) Myxovalargin, a long discovered myxobacterial compound, also contains a OH-Val residue,<sup>[9;10]</sup> however the biosynthetic pathway is only being currently investigated. (U. Scheid, personal communication)

To our knowledge, an enzyme which catalyzes the hydroxylation of valine at C $\beta$  is not yet known. The hydroxylation of Val in myxoprincomide possibly occurs as a post assembly or post incorporation modification by a yet to be identified oxidizing enzyme. We raise this possibility because the binding pocket of module 3 A domain (8 Å signature: LGSAFDASTLEGWLLAGGDVNGYGPTSTTFTAC / nonribosomal code: DALWLGGTFK) showed very high similarity to a Val-activating A domain. (Table S3.5) Both the 8 Å signature and the nonribosomal code lack an amino acid with a basic side chain (e.g. Arg or His) commonly observed in the binding pockets of hydroxyl group containing amino acid Ser- and Thr-activating A domains.<sup>[11]</sup>

Similar to the aminomutase MXAN\_4699, it was presumed that the necessary gene/s for the biosynthesis of OH-Val is found somewhere else on the genome as the genes surrounding *mxp*<sub>1622</sub> do not appear to encode proteins for hydroxylation. To facilitate the search, the amino acid for MPS1, a leucine hydroxylase which hydroxylates one of the C $\delta$  of leucine as the first step of (2S-4S) methylproline biosynthesis (Y. Katsuyama, unpublished data) was used as tblastn search query. The BLAST search analysis yielded three candidates: MXAN\_1289, MXAN\_4531, and MXAN\_3880. However, discrete targeted inactivation of all three candidate genes did not yield any myxoprincomide-negative mutants.

### 4.3 Tn5 promoter insertion upstream of NRPS module 1

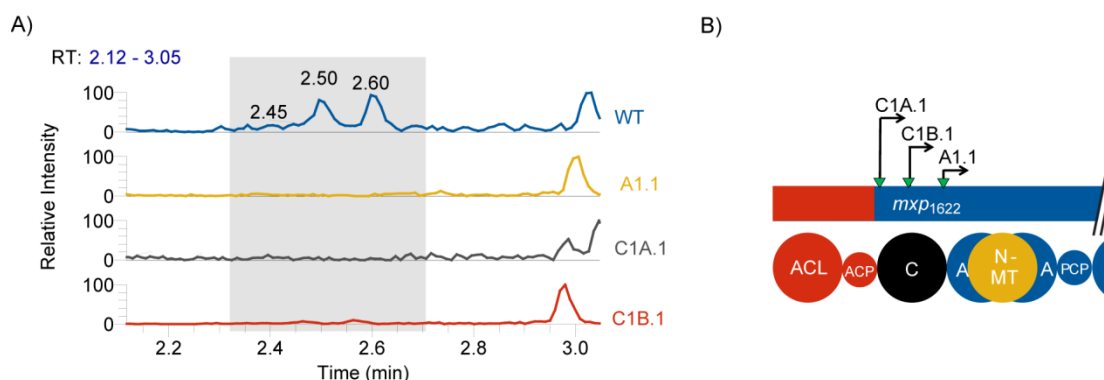
The N-terminus of Mxp contains a loading module comprised of an acyl-CoA ligase and ACP domains preceding the first NRPS module (module 1) which incorporates the N-MeSer residue. The C domain of module 1, in theory, is responsible for the condensation of an acyl moiety provided by the loading module and N-MeSer. However, since none



of the identified myxoprincomides contain an acyl moiety attached to the amine group of *N*-MeSer and that the module 1 C domain lack the complete set of required C domain motifs for activity (core motifs C5 and C7 are missing (Fig. S3.5)), this C domain is postulated to be inactive.

In order to ascertain whether the loading module is required, the Tn5 promoter sequence was inserted in frame upstream of the module 1 A domain gene sequence. Three mutant strains were generated (Fig. 4.22): one where the Tn5 promoter is directly upstream the 5' end of the module 1 A domain sequence (mutant strain A1.1, Tn5 promoter inserted upstream base 3,382) and the other two bear the Tn5 promoter in two distinct positions within the coding sequence of the module 1 C domain (mutant strain C1A.1, Tn5 promoter inserted upstream base 2,095; mutant strain C1B.1, Tn5 promoter inserted upstream base 2,860).

All three mutations resulted in a loss of myxoprincomide production. (Fig. 4.22) It is possible that the lack of the loading module and C domain sequence in the mutant strains' gene products have implications on the protein folding of the shortened Mxp<sub>1622</sub>, thereby resulting in a loss of biosynthetic activity.



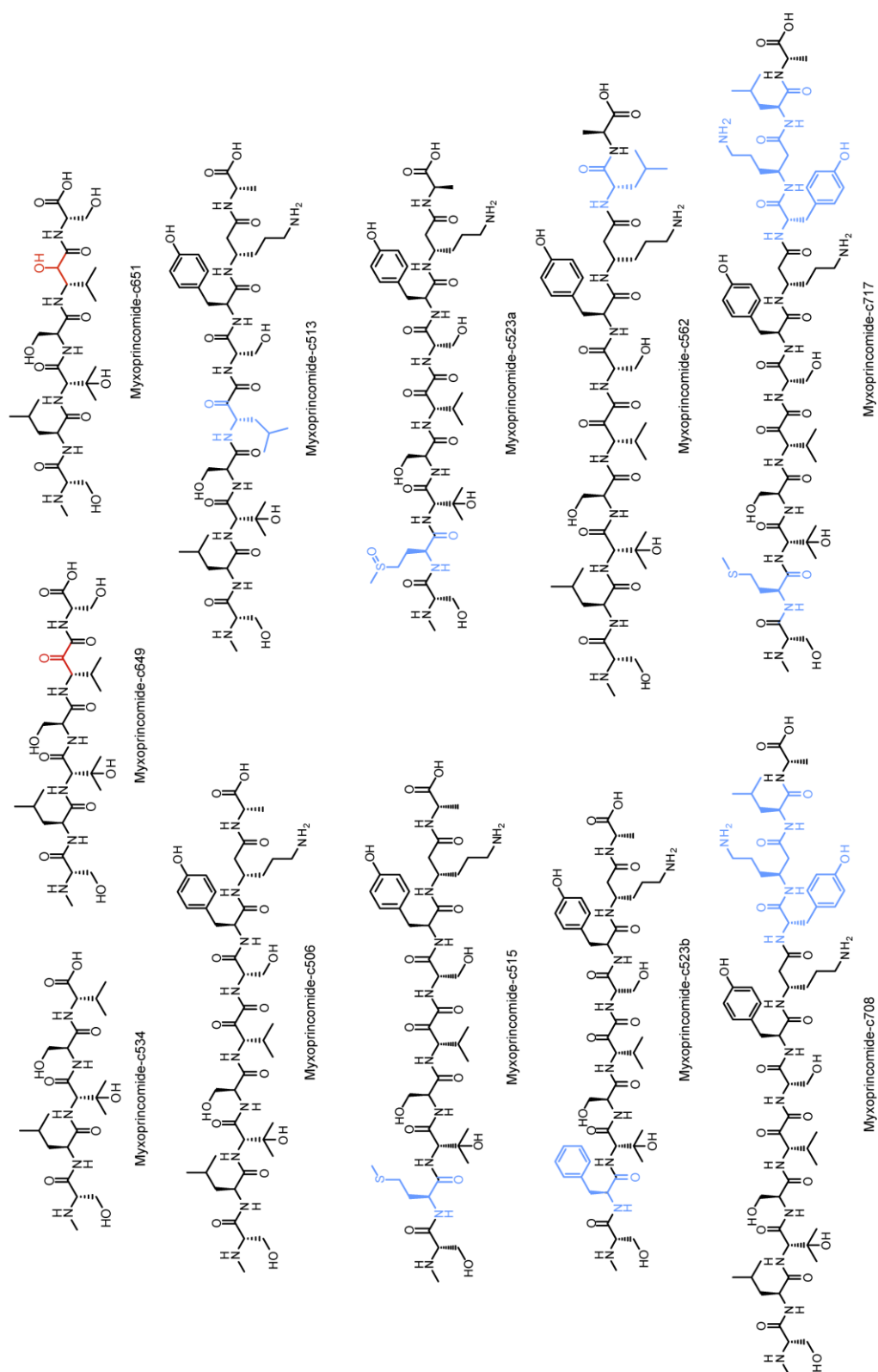
**Figure 4.22:** A) Extracted ion chromatograms of myxoprincomide from the wildtype and Tn5 promoter insertion mutants A1.1, C1A.1, and C1B.1. B) Arrows indicate the position where the Tn5 promoter was inserted in the *mxp1622* gene.

Alternatively, assuming that the module 1 C domain is active, it is possible that an acylated myxoprincomide is biosynthesized early during the growth phase but cleaved off at a later time point similar to the conversion of didemnin X and Y to didemnin B by *Tistrella mobilis* or the prodrug activation of xenocoumacin by the membrane-bound peptidase XcnG in *Xenorhabdus nematophila*.<sup>[12;13]</sup> However, XcnG does not cleave an acyl group per se but an amino acid (L-glutamine) to which the N-terminal acyl group is

bound to. A gene encoding a peptidase that could carry out this cleavage was not identified from the genes surrounding *mvp*. The closest candidate is the gene product of MXAN\_3780, but inactivation of MXAN\_3780 appears to have no effect on myxoprincomide biosynthesis. (see Section 4.4.1) Genes encoding proteins with sequence similarity to XcnG (Accession no. YP\_003711952, length 489 a.a.) were identified in the DK1622 genome (MXAN\_2364: length = 469 a.a., 27% identity, 48% similarity; MXAN\_6409: length: 524 a.a., 23% identity, 42% similarity; MXAN\_7171: length 741 a.a., 23% identity, 42% similarity). However, the involvement of the products of the aforementioned genes in myxoprincomide biosynthesis was not investigated in this study.

### 4.4 Unusual features of the group I Mxp<sub>1622</sub> assembly line

The structural diversity of the myxoprincomides encompass displaced monomeric units and derivatives of various peptide lengths. (Fig. 4.23) Based on domain organization, Mxp<sub>1622</sub> is predicted to assemble a lipopeptide comprised of 10 amino acid residues, an acyl chain at the N-terminus, and a malonate-derived C<sub>2</sub> unit in the middle of the peptide. Myxoprincomide-c506 could be presumed as the primary Mxp<sub>1622</sub> product given its relatively higher abundance, leading to its early detection prior to any of the other myxoprincomide variants. In itself, myxoprincomide-c506 deviates from strict colinearity by lacking the N-terminal acyl chain, the L-leucine residue near the C-terminus, and a carbon from the incorporated C<sub>2</sub> unit. The discovery of a multitude of other myxoprincomide variants demonstrates the atypical nature of Mxp<sub>1622</sub> for linear nonribosomal peptide biosynthesis.



**Figure 4.23:** Members of the group I myxoprincomides. Atoms and bonds in red represent the difference between myxoprincomides-c649 and -c651. Atoms and bonds in blue highlight the difference of the myxoprincomide variants from the representative structure myxoprincomide-c506

### 4.4.1 Release of intermediates

The identification of 5-er (myxoprincomide-c534) and 6-er (myxoprincomide-c649 & myxoprincomide-c651) shunt products highlights one of the unusual features of the myxoprincomide assembly. (Fig. 4.25)

Assembly line efficiency is dependent on seamless substrate channeling within the modules of the enzyme.<sup>[14]</sup> For example, the interaction of a C domain and the upstream and downstream PCP domains dictate the timing of the peptidyl chain growth.<sup>[15]</sup> Any inter- or intramolecular kinetic blocks that may delay or hamper substrate channeling therefore affects the overall biosynthetic turnover. Therefore, release of shunt products could be viewed as an adaptive mechanism, unfavorable in terms of the overall biosynthetic efficiency of the full product. However, the release of intermediates does expand the chemical diversity of the compound family.

The lack of 1- to 4-er and 8-er intermediates from the crude extracts insinuates that the release of the intermediates is not simply a matter of non-enzymatic hydrolysis based on the labile nature of the peptidyl-thioester bond.<sup>[16]</sup> The production 5-er (myxoprincomide-c534) and 6-er (myxoprincomide-c649 and -c651) are pronounced in large scale fermentation enabling their detection, isolation, and structural elucidation signifying that the corresponding modules from which they are released are hotspots for premature hydrolysis. (Fig. 4.25) A mass for the 7-er intermediate ( $m/z$   $[M+H]^+ = 812.4040$  Da, predicted formula  $C_{36}H_{58}N_7O_{14}$ ,  $\Delta ppm = 0.448$ ) was detected from the crude methanol extract from the large scale cultivation of *M. xanthus* A2.c506. Unfortunately, this putative intermediate neither endured the isolation work up nor was detected from the crude extracts from 50-ml scale cultivation, therefore structural elucidation by NMR or de novo peptide sequencing by tandem MS was not possible.

C domains which are unable or weakly able to recognize amino acids in their acceptor sites are hypothesized to release the intermediate when a water molecule enters the C domain acceptor site and reacts with the upstream peptidyl-thioester.<sup>[15]</sup> Expanding on this idea, a C domain or a KS domain, which is “waiting” on the downstream PCP domain to provide a substrate for the acceptor site, could also possibly facilitate a hydrolytic release of the intermediates. Following this hypothesis, it is possible that the group I A myxoprincomides are stalled peptidyl intermediates released as a consequence of the pressure of the next round of assembly. Formation of the  $\alpha$ -keto

moiety could be the first rate determining step during the biosynthesis and instead of stalling the assembly, myxoprincomide-c534 might be released. In the case of myxoprincomide-c649 and -c651, module 9 is occupied with the second round of L-tyrosine extension required for the biosynthesis of myxoprincomide-c708. The shorter nascent chain is then stalled at module 8 and is released by hydrolysis. Although there are type II thioesterases which act as proofreading enzymes in NRPS pathways, these TELLs act on aminoacyl substrates and not on peptidyl substrates.<sup>[17]</sup> C domain-mediated release of peptidyl intermediates is the current proposal for the formation of NRP shunt products.<sup>[16]</sup> However, the exact mechanism is not yet fully understood.

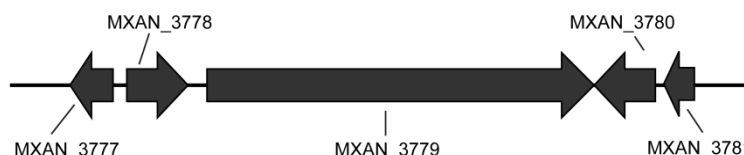
Hydrolytic release of upstream NRP intermediates were observed from the calcium-dependent antibiotic biosynthetic pathways after the A domains were engineered by site directed mutagenesis to alter specificity. It was proposed that the acceptor site of the C domain did not recognize the noncognate amino acid substrate thereby resulting in kinetic blockage.<sup>[16]</sup>

The release of shunt products by hydrolysis was previously observed in two PKS pathways when the genes encoding the release mechanisms of bacillaene (BaeR TE domain) and rifamycin (amide synthase RifF) were deleted.<sup>[18;19]</sup> The lack of an off-loading functionality resulted in the saturation of the assembly line with tethered linear intermediates. In the case of rifamycin, RifR, a type II thioesterase, was postulated to perform the off-loading of the intermediates from the assembly line. The release of the bacillaene intermediates was hypothesized to be carried out by an acyltransferase-like enzyme similar to PedC of the pederin biosynthesis.<sup>[20]</sup>

MXAN\_3780, a gene which putatively encodes a patatin-like phospholipase (PLP),<sup>[21]</sup> was initially thought to be involved in the release of the myxoprincomide intermediates. The predicted lipolytic activity of MXAN\_3780 is the closest biochemical functionality that could explain the release of shunt products. Inactivation of this gene however did not abrogate group I-A myxoprincomide production. Genes encoding PLPs are also found in the DKxanthene, leupyrrin, and tubulysin biosynthetic gene clusters<sup>[22-24]</sup> Inactivation of the PLP-encoding genes *orf17* and *orf18* of the tubulysin pathway led to the significant reduction of the production of the major tubulysins but led to the discovery of tubulysin-glycerol esters, deeming the PLPs responsible for the cleavage of

glycerol from the C-terminus of the peptide.<sup>[25]</sup> The exact role of PLPs in DKxanthene, leupyrrin, and now myxoprincomide biosyntheses has not been established.

Gene disruption of MXAN\_3778, which encodes a molecular chaperon DnaK, also did not affect the biosynthesis, not even a polar effect on the downstream *mxp*<sub>1622</sub> gene whose start codon is 143 bp away from the stop codon of MXAN\_3778. (Fig. 4.24)



**Figure 4.24:** Genetic environment of *mxp*<sub>1622</sub> (MXAN\_3779). The predicted functions for the gene products are as follows: MXAN\_3777 - Inosine-5'-monophosphate dehydrogenase; MXAN\_3778 - DnaK family protein; MXAN\_3779 - hybrid PKS / NRPS; MXAN\_3780 - Patatin-like phospholipase family protein; and MXAN\_3781 - DedA family protein. (See Table S3.4 for the source of homologous proteins)

With the exception of the PedC-mediated release of bacillaene intermediates, most of the previous observations of shunt product release were results of artificial conditions wherein the release mechanisms or the A domains were inactivated or modified. To our knowledge there has been no previous report of NRP shunt products that are released from the natural assembly line making myxoprincomide-c534, -c651, and c649 the first clear example of this phenomenon in NRP assembly.

### 4.4.2 Sluggish NRPS module and NRPS module stuttering

In the biosynthetic hypothesis of myxoprincomide-c506, we initially proposed that module 11 incorporates the final Ala monomer into the nascent peptide chain and is released by the TE domain, skipping the C-A-PCP of module 12 of *Mxp*<sub>1622</sub>.<sup>[26]</sup> However the discovery of myxoprincomide-c562 and myxoprincomide-c708 provided evidence which forced us to revise our previous biosynthetic hypothesis. (Section 3.4) In this revision, we propose that module 12 is active and is responsible for the Ala extension and that module 11 is a sluggish NRPS module which sometimes elongates the peptide chain with a Leu.

We describe module 11 as a sluggish NRPS module because majority of the large myxoprincomides do not contain Leu in the penultimate position, thereby suggesting that the module is skipped. Nevertheless, we cannot omit the possibility that the

module is active as all the conserved motifs required for catalysis are present in the C, A, and PCP domains of module 11. (Fig. S3.5) This is unlike in the skipping of the NRPS module 4 in the myxochromide biosynthesis where core motifs of the A4 domain significantly deviates from conserved A domain motifs required for activity and PCP4 lacks the catalytic serine residue.<sup>[27]</sup>

A possible scenario is that the activation and loading of Leu at module 11 occurs, resulting in myxoprincomide-c562, but the process is perhaps not efficient wherein there is no Leu loaded onto module 11 thereby stalling the biosynthesis at module 10. Instead of releasing the stalled peptide similar to the group I-A myxoprincomides, the Mxp<sub>1622</sub> synthetase alternatively allows the peptidyl-thioester at module 10 to condense with the PCP-bound amino acids of a) module 11 resulting in module skipping forming myxoprincomide-c506, -c515, and the group I-C myxoprincomides, or b) module 9 exhibiting NRPS stuttering where the nascent chain at module 10 condenses with another L-tyrosine bound to the PCP domain module 9. The chain elongates further with a second pass through module 10 for a second L- $\beta$ -lysine extension followed by the condensation with L-leucine and L-alanine in modules 11 and 12 resulting in myxoprincomide-c708. The occurrence of NRPS isomodular stuttering, or usage of the same module multiple times, is known for the biosynthesis of saframycin A and congocidine from streptomycetes.<sup>[28;29]</sup> Iteration in NRPS where extension with precursors tethered to the modules upstream of the immediate elongation module, termed here as retromodular stuttering, has been recently reported in the biosynthesis of thalassospiramides from marine bacteria.<sup>[30]</sup> Modules 2 to 4 of the NRPS assembly line undergo a second, and even a third, round of extension to expand the thalassospiramide chemical diversity. The exact mechanism is yet to be fully understood but it is hypothesized that the PCP-PCP didomain in module 4 may be involved in the substrate shuttling in order for the iterative assembly to occur. In the case of Mxp<sub>1622</sub>, no such PCP didomain is present in any of the modules. The only “free” thiolation (ACP) domain is that of the loading module. However, the involvement of the loading module ACP domain is speculative, as its role in myxoprincomide biosynthesis is not yet fully understood.

#### 4.4.3 Substrate promiscuity of the A domains of Mxp<sub>1622</sub> module 2 and module 6

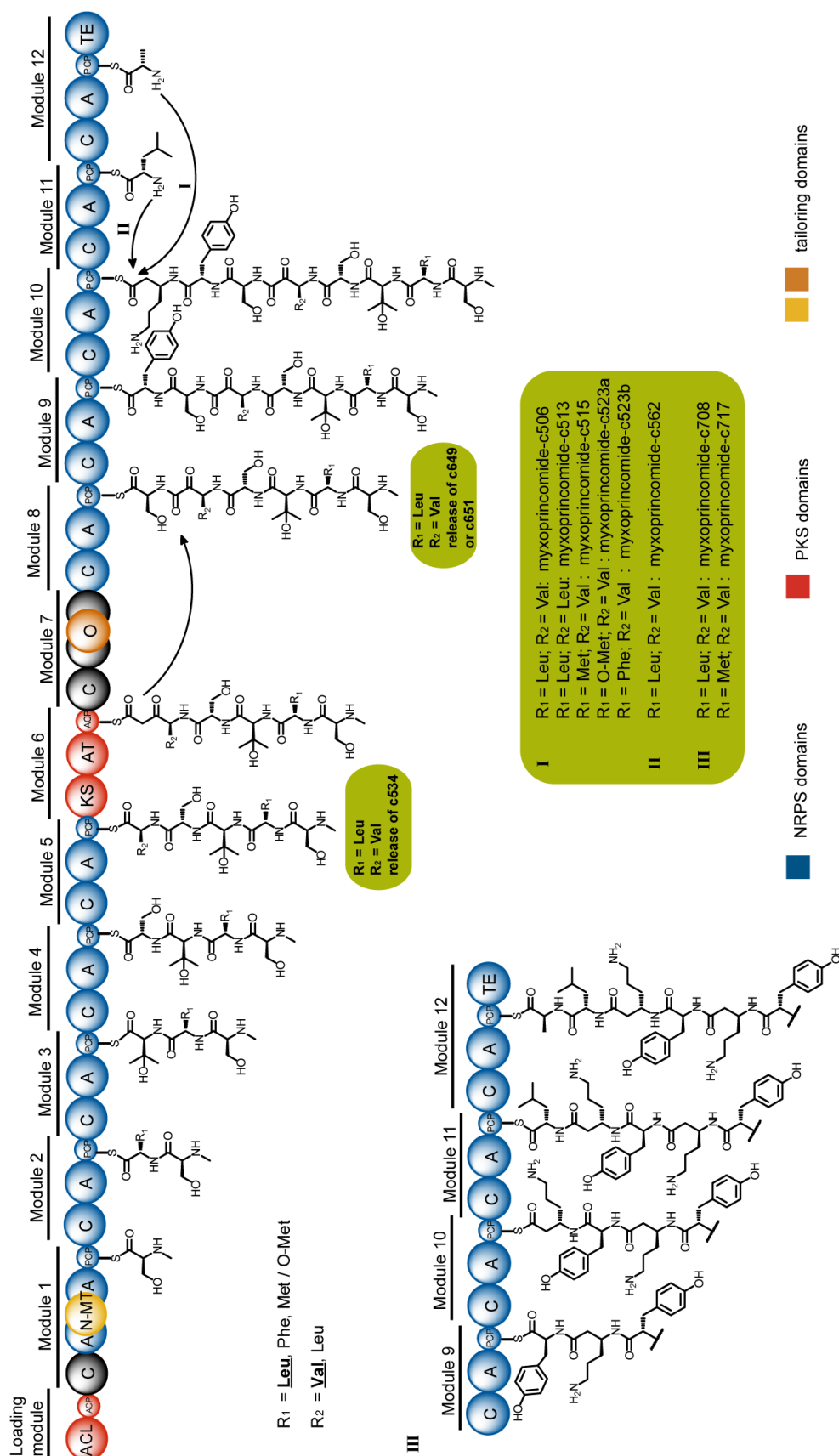
The group I-C myxoprincomides are distinct from the other groups in that the amino acid at position 2 of the molecules deviates from the more commonly observed Leu

residue. The incorporation of methionine (myxoprincomide-c515 and -c523a (as methionine sulfoxide in the latter)) or phenylalanine suggests substrate promiscuity in the A domain of module 2. Coincidentally, the predicted nearest neighbor for the module 2 A domain based on *in silico* analysis is the TycB module 2 A domain of the tyrocidine biosynthetic pathway, which itself flexibly recognizes both phenylalanine and tryptophan.<sup>[31]</sup>

The observation that Leu is the preferred substrate and that Met and Phe are alternative substrates suggests that the Mxp<sub>1622</sub> module 2 A domain is amenable to amino acids with hydrophobic side chains. The observation of the methionine-S-oxide containing variant myxoprincomide-c523a is somewhat anomalous as the sulfoxide moiety is polar. This residue is observed in known nonribosomal peptides from marine sponge (haligramide B & waiakemide) and cyanobacteria (nodulapeptin B, somamides).<sup>[32-35]</sup> It is therefore proposed that the formation of the sulfoxide occurs as a post-incorporation nonenzymatic oxidation, similar to the proposed biosynthesis of nodulapeptin C.<sup>[34]</sup> However, we could also not exclude the possibility that methionine sulfoxide is incorporated as the binding pocket represented by the 8 Å signature (LGTSFDVSSVEGLLMVGGETNAYGPTECSTMVTR) contains a number of amino acids with polar side chains that could interact with the polar sulfoxide functional group. Establishing the molecular recognition profile of the Mxp<sub>1622</sub> module 2 A domain requires experimental proof by ATP-pyrophosphate exchange assay *in vitro* using recombinant A domain similar to that used for TycA.<sup>[36]</sup>

The substrate plasticity of module 6 A domain is simpler in its range of substrates. Majority of the products observed contains the Val residue west of the  $\alpha$ -keto moiety. Leucine is accepted as well (myxoprincomide-c513) but to a lesser extent. The additional methylene group in Leu likely lowers the catalytic efficiency of the module 6 A domain, resulting in the low production rate of myxoprincomide-c513. The same trend is observed in the group II myxoprincomides produced by *M. xanthus* DK897 where the intensity of the Val-containing variants is more than twice than that of the Leu-containing group II myxoprincomides (Fig. S4.25). (O. Revermann, Dissertation, Universität des Saarlandes, 2012)





**Figure 4.25:** Biosynthetic model for the assembly of myxoprincomides. R1 and R2 denote the residues where multiple monomers are possible. Underlined is the most commonly observed monomers for R1 and R2. I – pathway which skips module 11. II – pathway which includes a Leu elongation facilitated by module 11. III – retromodular stuttering with a second round of Tyr and β-Lys elongation followed by a Leu and Ala extension by modules 11 and 12. In green are the released myxoprincomide products

### 4.4.4 Revised and extended model for the biosynthesis of myxoprincomides

Based on the spectrum of myxoprincomides produced by *M. xanthus* DK1622, we revise the previously presented biosynthetic model to accommodate the structural features observed in the products. (Fig. 4.25)

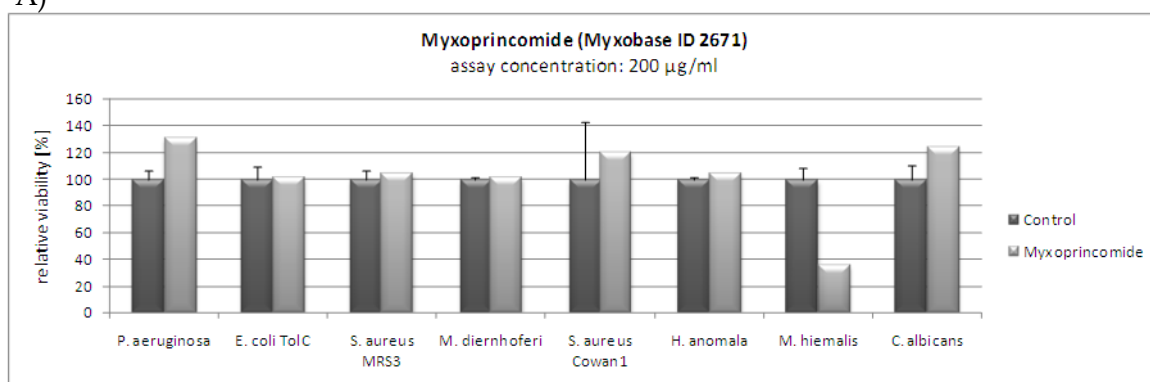
Since none of the myxoprincomide variants identified so far contain an acyl starter unit, we will presume that condensation between a hypothetical acyl chain on the loading module and the L-serine precursor does not occur due to an inactive C domain in module 1. Therefore biosynthesis is hypothesized to begin in module 1 with the incorporation and *N*-methylation of L-serine. The substrate promiscuity of the module 2 allows the catalytic flexibility to select, activate, and load L-leucine, L-methionine (and possibly L-methionine sulfoxide), or L-phenylalanine, with L-leucine as the preferred substrate. Chain extension proceeds with the incorporation of L- $\beta$ -OH-valine (possibly a L-valine first followed by a hydroxylation event) and L-serine by modules 3 and 4. Module 5 extends the peptide with L-valine but the module also accepts a L-leucine (myxoprincomide-c513) albeit at a lower catalytic efficiency. At this stage, the module 5 PCP-tethered intermediate could either be released forming myxoprincomide-c534 or extension could occur with the downstream PKS module. Similar to our previous biosynthetic model, we propose that the PKS module 6 elongates the nascent chain with a malonyl-CoA followed by an unusual oxidation of the carbon at position  $\alpha$  to the thioester function presumably catalyzed by the module 7 oxidation domain. (Fig. S3.17) Tautomerization possibly forms an  $\alpha$ -keto moiety which is the preferred nucleophilic site of the downstream L-serine amino group. We conclude this based on the fact that feeding with  $^{13}\text{C}$ -labeled acetate consistently maintained only the carbon at position 2 in all of the  $\alpha$ -keto-containing myxoprincomides. (Table S4.1) The L-serine extended peptide could either prematurely hydrolyze to release myxoprincomide-c649 or myxoprincomide-c651 or chain extension proceeds with the incorporation of L-tyrosine and L- $\beta$ -lysine by modules 9 and 10 respectively. The next elongation step has three possible scenarios. The most common based on the observed product profile is the skipping of module 11 and the consequent elongation with an L-alanine precursor by module 12 followed by a TE-facilitated release. The second possibility is that module 11 is used and a L-leucine is incorporated before the L-alanine extension and is released as myxoprincomide-c562 by the TE domain. The last route is the presumed retromodular stuttering wherein the module 10 PCP-bound intermediate is extended by a repeat

extension with L-tyrosine and L- $\beta$ -lysine by modules 9 and 10 and the consequent incorporation of L-leucine and L-alanine by modules 11 and 12 to form myxoprincomide-c708.

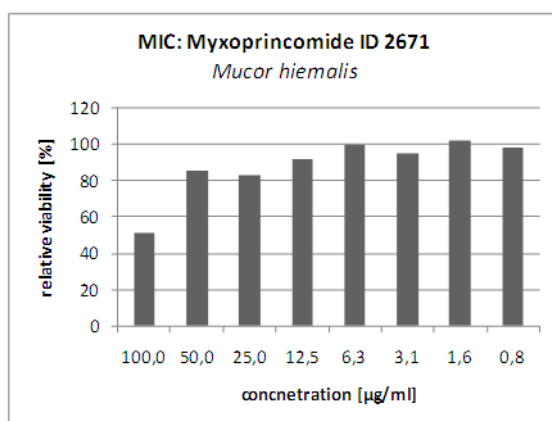
## 4.5 Bioactivity

Myxoprincomide-c506 and myxoprincomide-c708 underwent extensive activity screening spanning cytotoxicity, antibacterial, and antifungal assays. (Section 2.23.2) Myxoprincomide-c506 was not found to be active in any of the assays performed even up to a concentration of 200  $\mu\text{g mL}^{-1}$ . On the other hand, myxoprincomide-c708 was observed to weakly inhibit the growth of the filamentous fungi *Mucor hiemalis* at 200  $\mu\text{g mL}^{-1}$ . The minimum inhibitory concentration (MIC) of myxoprincomide-c708 was approximated to be between 50 – 100  $\mu\text{g mL}^{-1}$ . (Fig. 4.26)

A)



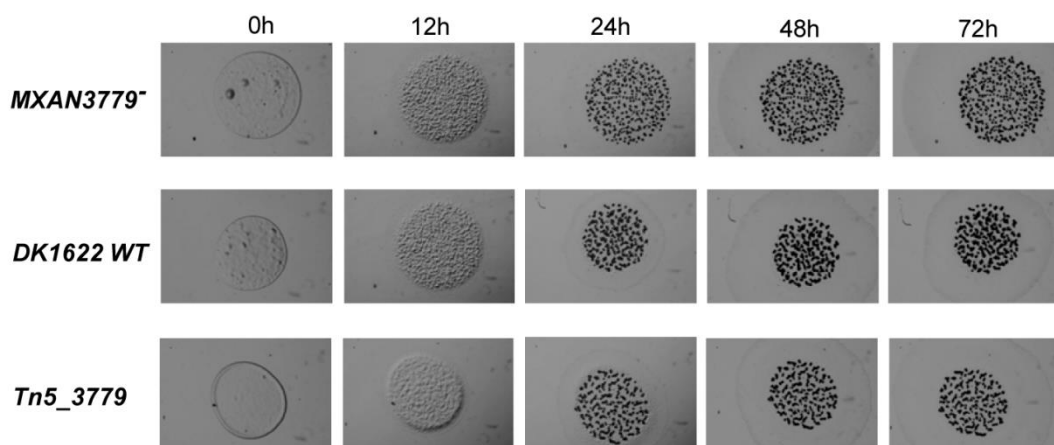
B)



**Figure 4.26:** Antimicrobial assay of myxoprincomide-c708 at 200  $\mu\text{g mL}^{-1}$  (A) and determination of MIC against *M. hiemalis* (B).

## 4 Myxoprincomides

To determine whether myxoprincomides are relevant for *M. xanthus* social development and motility, the DK1622 wildtype, *mxp*<sub>1622</sub> knockout mutant (MXAN\_3779-), and overproducer (Tn5\_3779) strains were analyzed by fruiting body development, sporulation frequency, and motility assays. Comparison of the formation of fruiting bodies over a course of three days did not reveal any significant difference among all three strains. (Fig. 4.27)



**Figure 4.27:** Fruiting body formation assay on MXAN\_3779-, wildtype, and Tn5\_3779 strains. Images were taken at the indicated time points above each column.

Formation of spores was also not affected by the inactivation or overexpression of *mxp*<sub>1622</sub> as viable colonies of comparable numbers were observed on agar plates after disruption and overlay of the formed spores. Swarming motility was not significantly impeded nor enhanced in the knockout and overproducing mutant strains as swarming was observed at approximately the same speed within the experimental error window. (Table 4.5)

**Table 4.5:** Average swarming speeds (mm/day) of the *M. xanthus* DK1622 wildtype, knockout mutant, and overproducer strains

	CTT Agar	CTT Soft Agar
Strain	Average mm / day	Average mm / day
DK1622	2.78	3.22
MXAN3779-	3.08	3.42
Tn5_3779	3.01	3.74

## 4.5 Outlook

Majority of the myxoprincomide structures presented in this chapter were structurally elucidated by tandem MS and LC-MS/MS<sup>2</sup> analysis of isotope-labeled myxoprincomides. Although the interpretation is sound, the proposed structures could be further confirmed by purification and subsequent analysis by NMR. Obtaining sufficient material for structural elucidation will be challenging but by scaling up the VY/2 fed batch cultivation, the availability of the myxoprincomides could be achieved. Moreover, increasing the number of purified variants affords more chemical entities for bioactivity assays and will enable to determine whether the variation in the structures present any evolutionary advantage for *M. xanthus*.

The molecular diversity of the myxoprincomides revealed the atypical nature of Mxp<sub>1622</sub>. The proposed biosynthetic model provided reflects a good agreement between the chemical reality and the genetic evidence. However, there are biosynthetic steps which still require *in vivo* or *in vitro* analysis for support.

An ATP-pyrophosphate exchange assay similar to that used for TycA<sup>[36]</sup> could provide empirical proof to establish the extent of the substrate promiscuity of Mxp<sub>1622</sub> module 2 A domain and additionally provide a kinetic evidence for the L-leucine incorporation bias. But the most interesting segment for further investigation would be modules 9 to 11 where retromodular stuttering and module skipping / usage occurs.

A deeper understanding of the PKS-derived formation of the  $\alpha$ -keto /  $\alpha$ -hydroxy moiety could be achieved by *in vitro* analysis using synthetic intermediates and a recombinant partial assembly line. Given the issues of expressing partial assembly lines, an alternative to fully elucidate the biosynthesis would be to construct truncated assembly lines with promoter insertion upstream of modules 6 and 7 separately and to feed with activated synthetic intermediates. The assumption mainly is that the truncated assembly line could still properly fold and is functional to carry out the rest of the biosynthesis. In addition, point mutations in the Ox domain could lead to domain inactivity possibly producing myxoprincomides without the  $\alpha$ -keto moiety.

Fusion of the TE domain upstream of the later modules forming a truncated Mxp<sub>1622</sub> could provide further insight into the release mechanism of the shunt products. This technique has been shown to provide truncated intermediates from polyketide synthase assembly lines.<sup>[18;37;38]</sup>

### 4.8 References

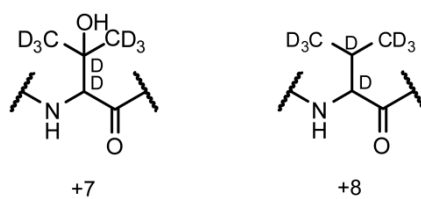
- [1.] D. Krug, G. Zurek, O. Revermann, M. Vos, G. J. Velicer, R. Müller, *Appl.Environ.Microbiol.* **2008**, 74 3058-3068.
- [2.] G. J. Velicer, Y. T. Yu, *Nature* **2003**, 425 75-78.
- [3.] L. Shimkets, M. Dworkin, H. Reichenbach, in *The Prokaryotes*, Vol. 7, (Ed.: M. Dworkin), Springer, Berlin **2006**, p. pp. 31-115.
- [4.] B. Wu, W. Szymanski, M. M. Heberling, B. L. Feringa, D. B. Janssen, *Trends Biotechnol.* **2011**, 29 352-362.
- [5.] J. J. Barkei, B. M. Kevany, E. A. Felnagle, M. G. Thomas, *ChemBioChem* **2009**, 10 366-376.
- [6.] N. Grammel, K. Pankevych, J. Demydchuk, K. Lambrecht, H. P. Saluz, U. Keller, H. Krugel, *Eur.J.Biochem.* **2002**, 269 347-357.
- [7.] B. W. Lepore, F. J. Ruzicka, P. A. Frey, D. Ringe, *P.Natl.Acad.Sci.USA* **2005**, 102 13819-13824.
- [8.] D. Pistorius, Y. Li, A. Sandmann, K. J. Weissman, R. Müller, *Molecular Biosystems* **2011**, 7 3308-3315.
- [9.] H. Irschik, K. Gerth, T. Kemmer, H. Steinmetz, H. Reichenbach, *J.Antibiot.* **1983**, 36 6-12.
- [10.] H. Steinmetz, H. Irschik, H. Reichenbach, G. Höfle, in *Chemistry of Peptides and Proteins - Proceedings of the Sixth UssR-FRG Symposium on Chemistry of Peptides and Proteins (Hamburg, FRG, Sept 1-5, 1987)* Eds.: W. A. König, W. Voelter), Attempto Verlag, Tübingen **1987**, p. pp. 13-18.
- [11.] M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch, O. Kohlbacher, *Nucleic Acids Res.* **2011**, 39 W362-W367.
- [12.] Y. Xu, R. D. Kersten, S. J. Nam, L. Lu, A. M. Al-Suwailem, H. Zheng, W. Fenical, P. C. Dorrestein, B. S. Moore, P. Y. Qian, *J.Am.Chem.Soc.* **2012**, 134 8625-8632.
- [13.] D. Reimer, K. M. Pos, M. Thines, P. Grüm, H. B. Bode, *Nat Chem Biol Nat Chem Biol* **2011**, 7 888-890.
- [14.] S. Y. Tsuji, N. Wu, C. Khosla, *Biochemistry-US* **2001**, 40 2317-2325.
- [15.] P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **1999**, 284 486-489.
- [16.] G. C. Uguru, C. Milne, M. Borg, F. Flett, C. P. Smith, J. Micklefield, *J.Am.Chem.Soc.* **2004**, 126 5032-5033.
- [17.] E. Yeh, R. M. Kohli, S. D. Bruner, C. T. Walsh, *ChemBioChem* **2004**, 5 1290-1293.
- [18.] J. Moldenhauer, X. H. Chen, R. Borriss, J. Piel, *Angew.Chem.Int.Ed.* **2007**, 46 8195-8197.

- [19.] T. W. Yu, Y. Shen, Y. Doi-Katayama, L. Tang, C. Park, B. S. Moore, C. R. Hutchinson, H. G. Floss, *P.Natl.Acad.Sci.USA* **1999**, 96 9051-9056.
- [20.] K. Jensen, H. Niederkrüger, K. Zimmermann, A. L. Vagstadt, J. Moldenhauer, N. Brendel, S. Frank, P. Pöplau, C. Kohlhaas, C. A. Townsend, M. Oldiges, C. Hertweck, J. Piel, *Chemistry & Biology* **2012**, 19 329-339.
- [21.] B. S. Goldman, W. C. Nierman, D. Kaiser, S. C. Slater, A. S. Durkin, J. Eisen, C. M. Ronning, W. B. Barbazuk, M. Blanchard, C. Field, C. Halling, G. Hinkle, O. Iartchuk, H. S. Kim, C. Mackenzie, R. Madupu, N. Miller, A. Shvartsbeyn, S. A. Sullivan, M. Vaudin, R. Wiegand, H. B. Kaplan, *P.Natl.Acad.Sci.USA* **2006**, 103 15200-15205.
- [22.] P. Meiser, K. J. Weissman, H. B. Bode, D. Krug, J. S. Dickschat, A. Sandmann, R. Müller, *Chem.Biol.* **2008**, 15 771-781.
- [23.] M. Kopp, H. Irschik, K. Gemperlein, K. Buntin, P. Meiser, K. J. Weissman, H. B. Bode, R. Müller, *Molecular Biosystems* **2011**, 7 1549-1563.
- [24.] Y. Chai, D. Pistorius, A. Ullrich, K. J. Weissman, U. Kazmaier, R. Müller, *Chem.Biol.* **2010**, 17 296-309.
- [25.] Chai, Yi, Universitäts- und Landesbibliothek, **2011**.
- [26.] N. S. Cortina, D. Krug, A. Plaza, O. Revermann, R. Müller, *Angew.Chem.Int.Ed.* **2012**, 51 811-816.
- [27.] S. C. Wenzel, B. Kunze, G. Höfle, B. Silakowski, M. Scharfe, H. Blöcker, R. Müller, *ChemBioChem* **2005**, 6 375-385.
- [28.] L. Li, W. Deng, J. Song, W. Ding, Q. F. Zhao, C. Peng, W. W. Song, G. L. Tang, W. Liu, *J.Bacteriol.* **2008**, 190 251-263.
- [29.] M. Juguet, S. Lautru, F. X. Francou, S. Nezbedova, P. Leblond, M. Gondry, J. L. Pernodet, *Chem Biol* **2009**, 16 421-431.
- [30.] A. C. Ross, Y. Xu, L. Lu, R. D. Kersten, Z. Shao, A. M. Al-Suwailem, P. C. Dorrestein, P. Y. Qian, B. S. Moore, *J.Am.Chem.Soc.* **2012**.
- [31.] H. D. Mootz, M. A. Marahiel, *J.Bacteriol.* **1997**, 179 6843-6850.
- [32.] M. A. Rashid, K. R. Gustafson, J. L. Boswell, M. R. Boyd, *J.Nat.Prod.* **2000**, 63 956-959.
- [33.] C. M. S. Mau, Y. Nakao, W. Y. Yoshida, P. J. Scheuer, M. KellyBorges, *J.Org.Chem.* **1996**, 61 6302-6304.
- [34.] L. Rouhiainen, J. Jokela, D. P. Fewer, M. Urmann, K. Sivonen, *Chem.Biol.* **2010**, 17 265-273.
- [35.] L. M. Nogle, R. T. Williamson, W. H. Gerwick, *J Nat.Prod.* **2001**, 64 716-719.
- [36.] B. R. Villiers, F. Hollfelder, *ChemBioChem* **2009**.

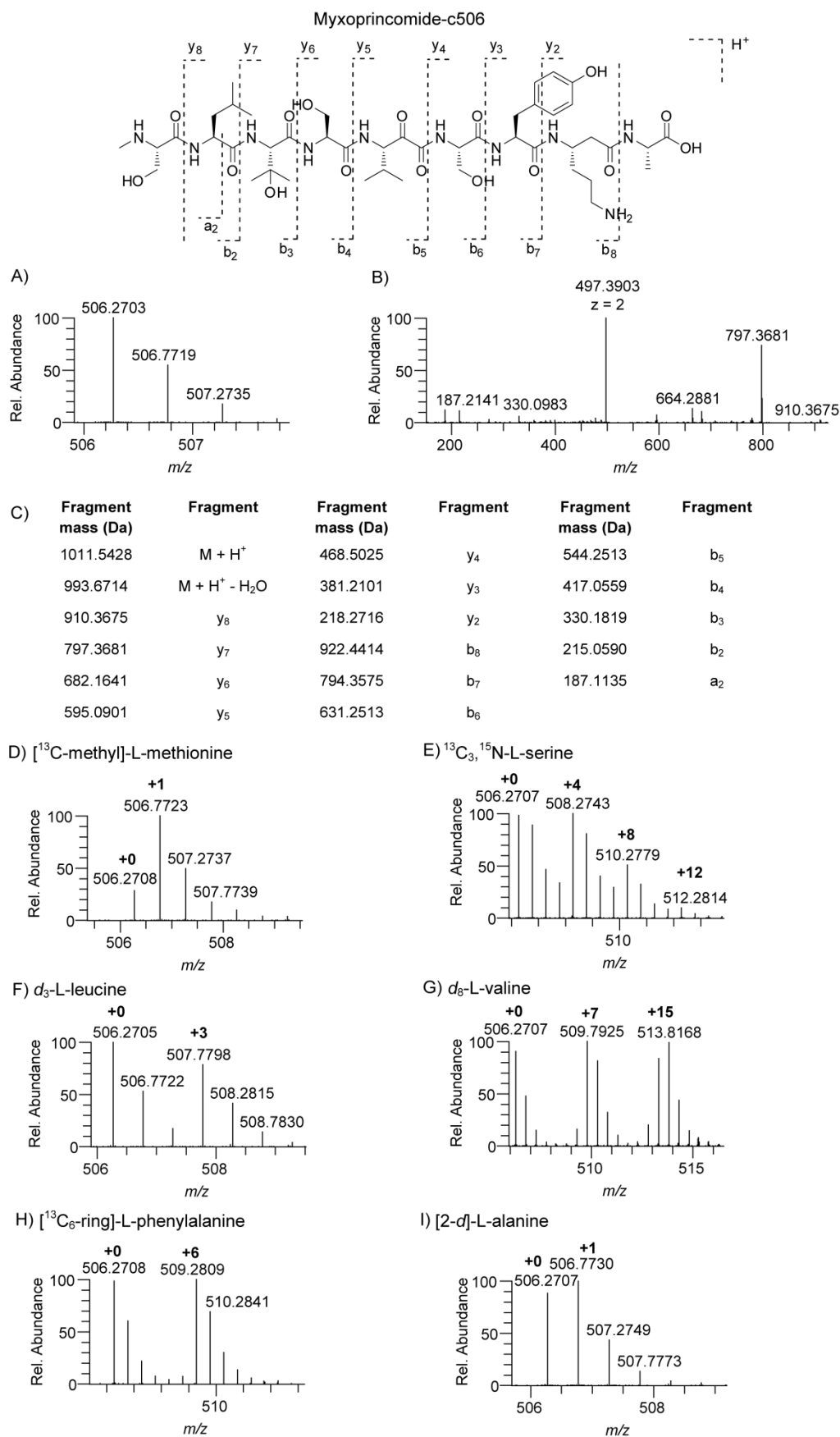
- [37.] C. J. Martin, M. C. Timoney, R. M. Sheridan, S. G. Kendrew, B. Wilkinson, J. C. Staunton, P. F. Leadlay, *Org.Biomol.Chem.* **2003**, *1* 4144-4147.
- [38.] J. Cortes, K. E. H. Wiesmann, G. A. Roberts, M. J. B. Brown, J. Staunton, P. F. Leadlay, *Science* **1995**, *268* 1487-1489.



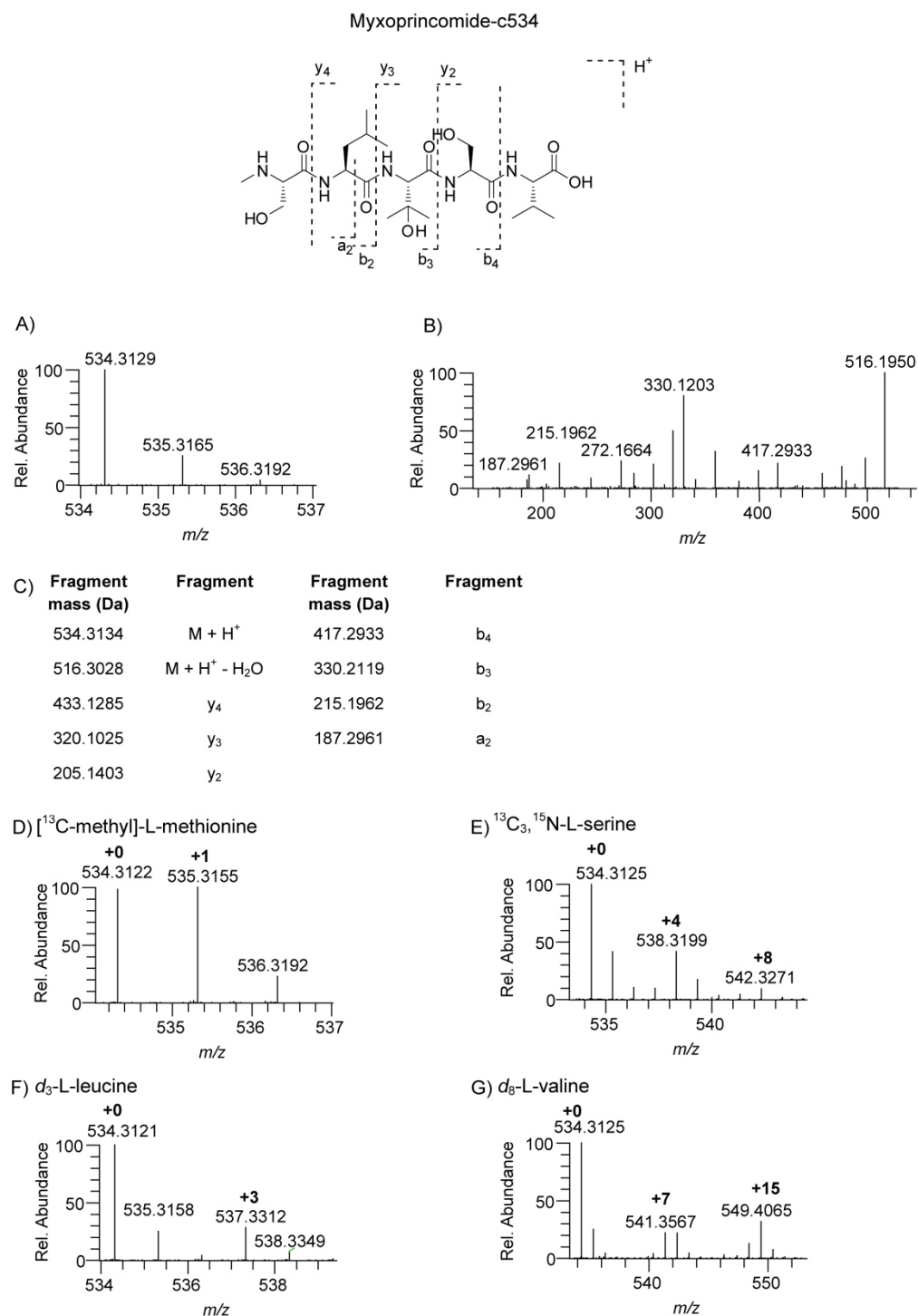
## 4.7 Supplementary Information



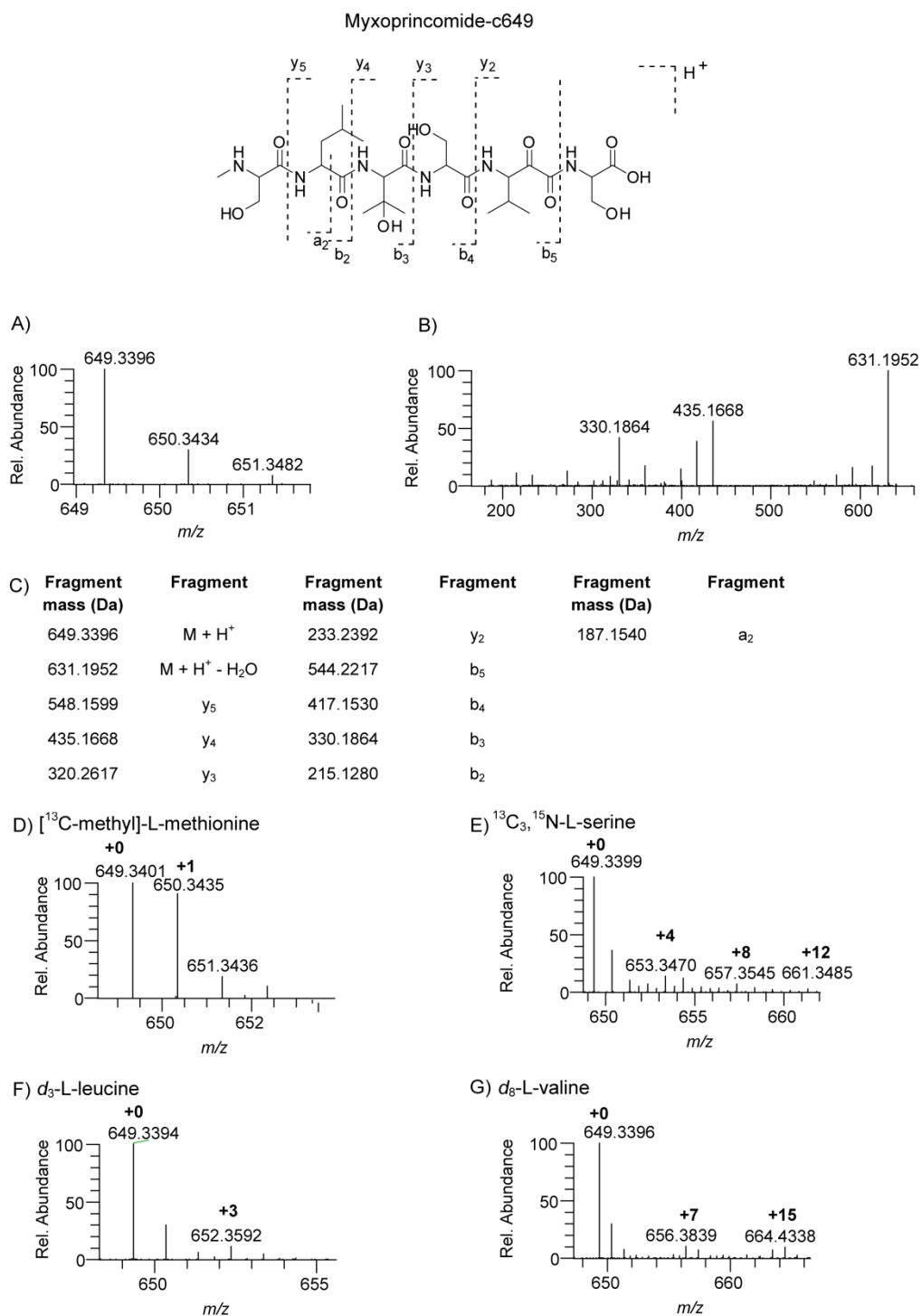
**Figure S4.1:** Deuterium-labeled L-β-OH-valine and L-valine residues.



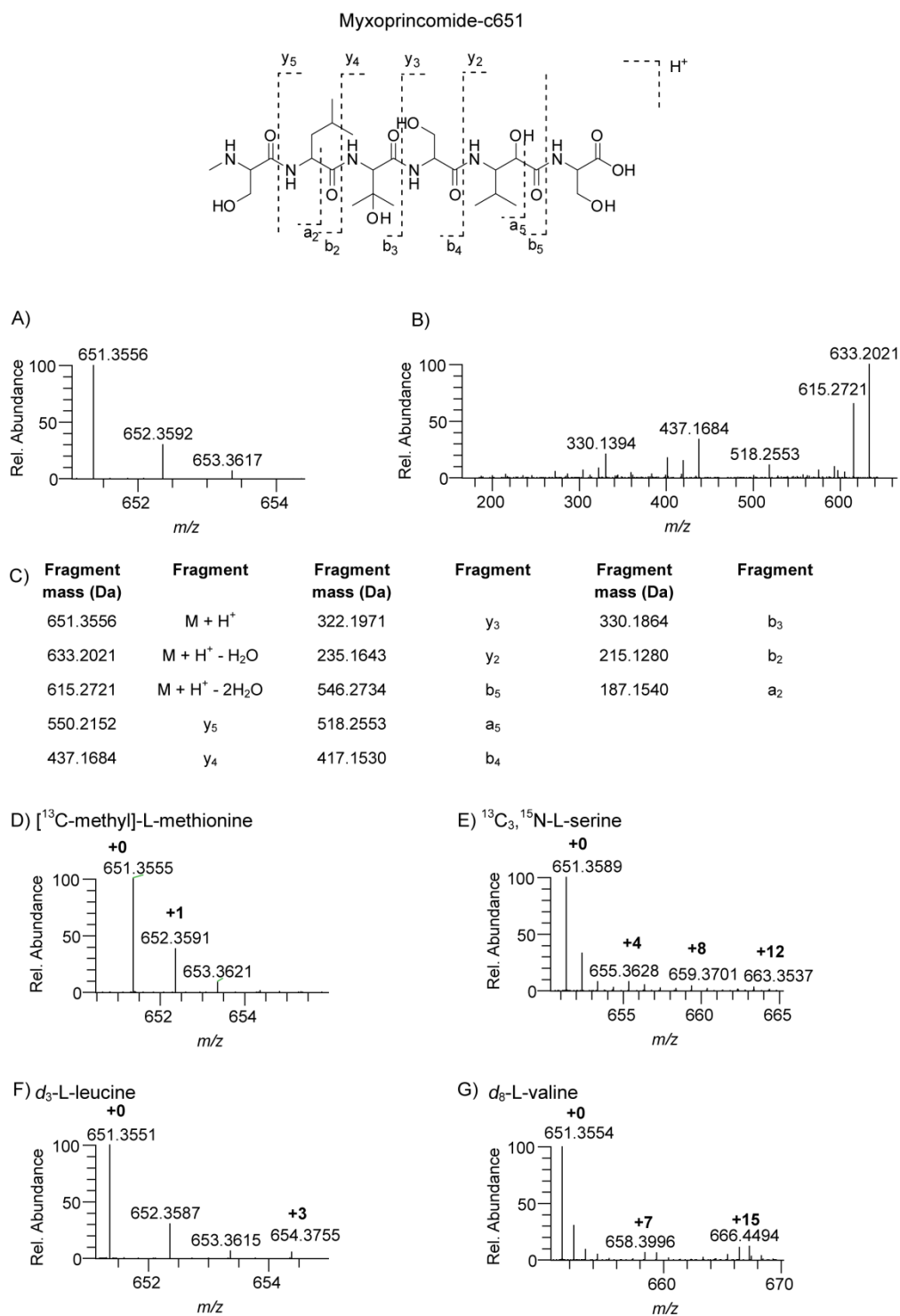
**Figure S4.2:** Myxoprincomide-c506 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-I) HR-ESI MS spectra of labeled compounds



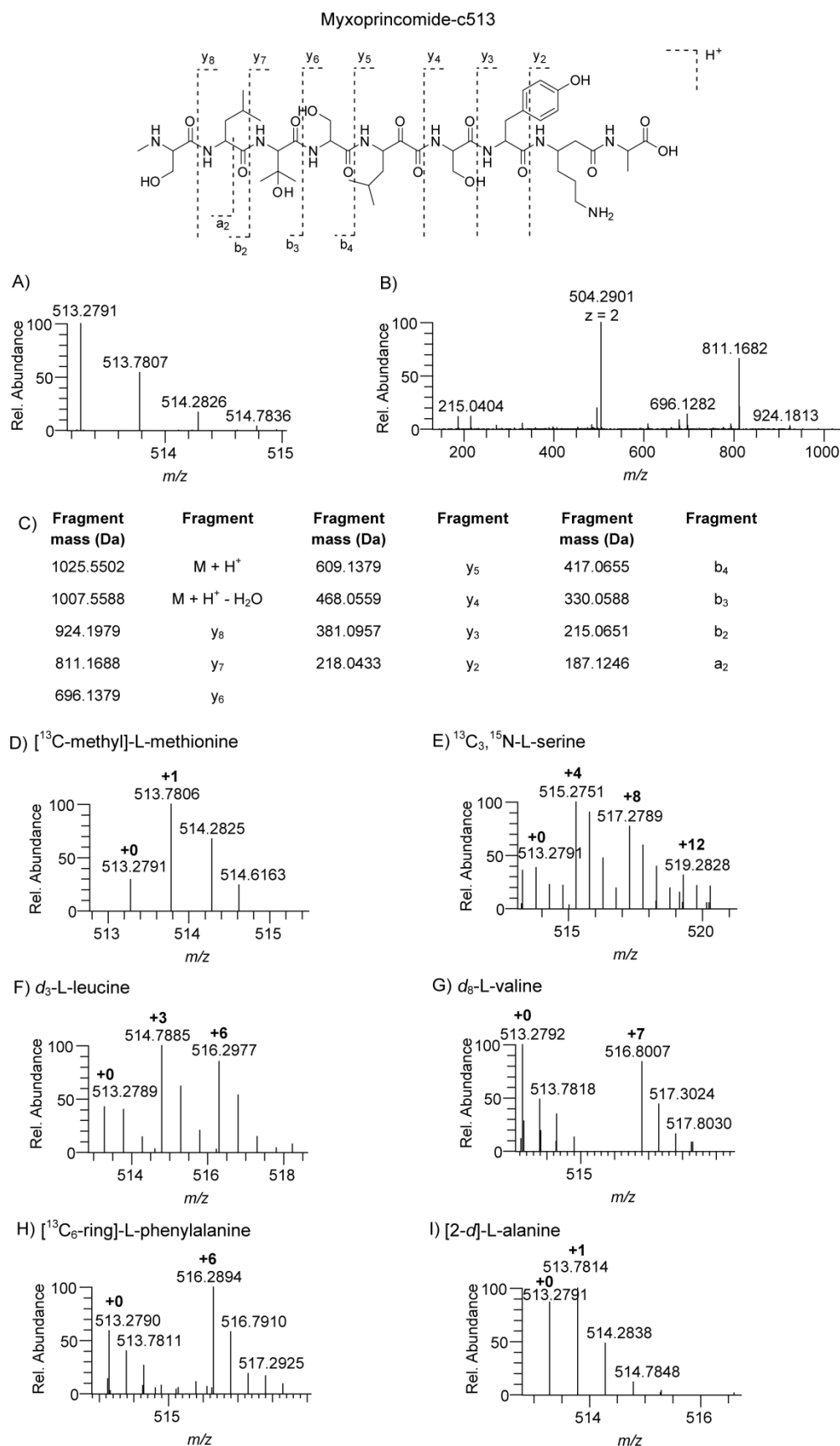
**Figure S4.3:** Myxoprincomide-c534 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds



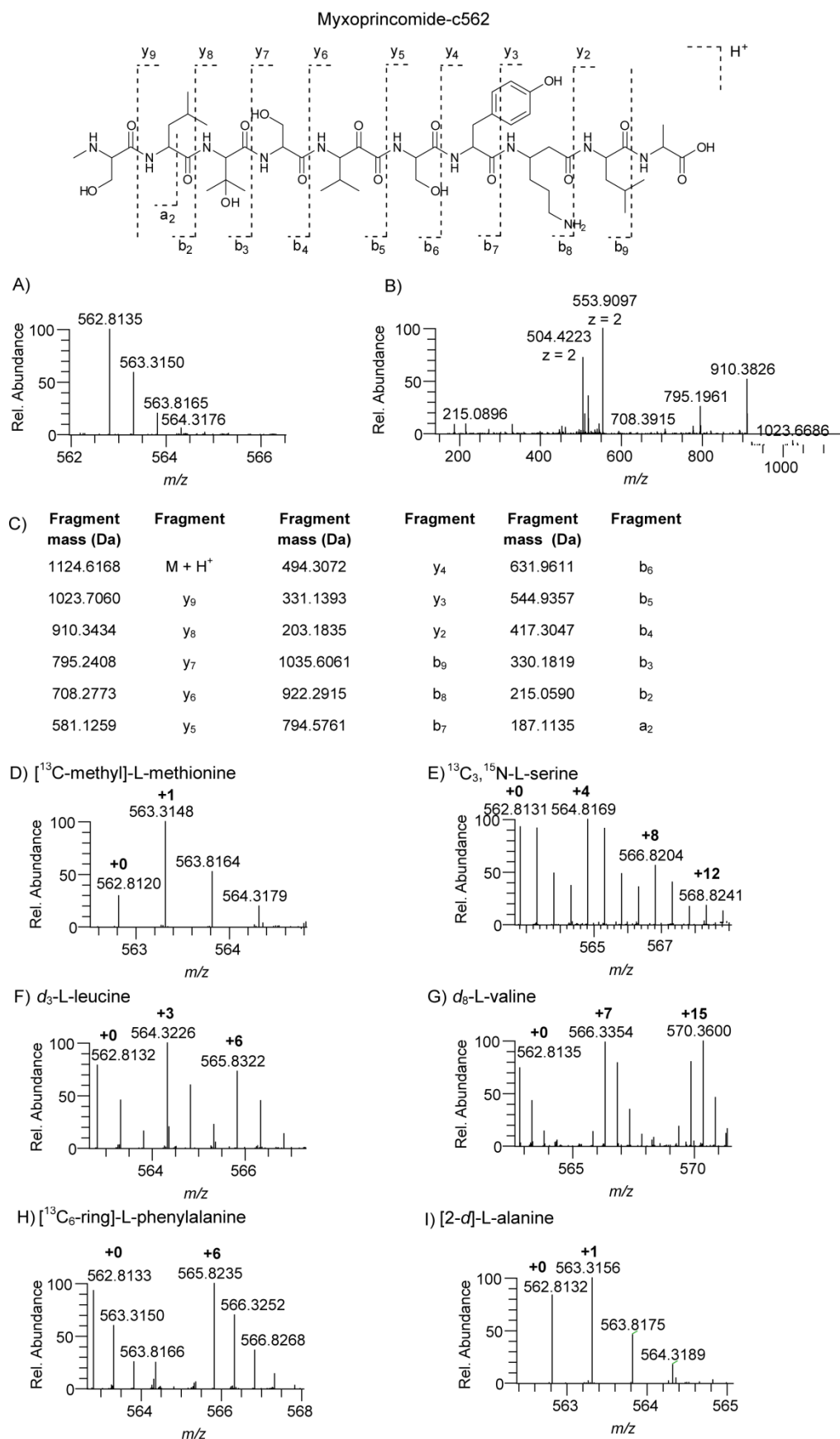
**Figure S4.4:** Myxoprincomide-c649 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds



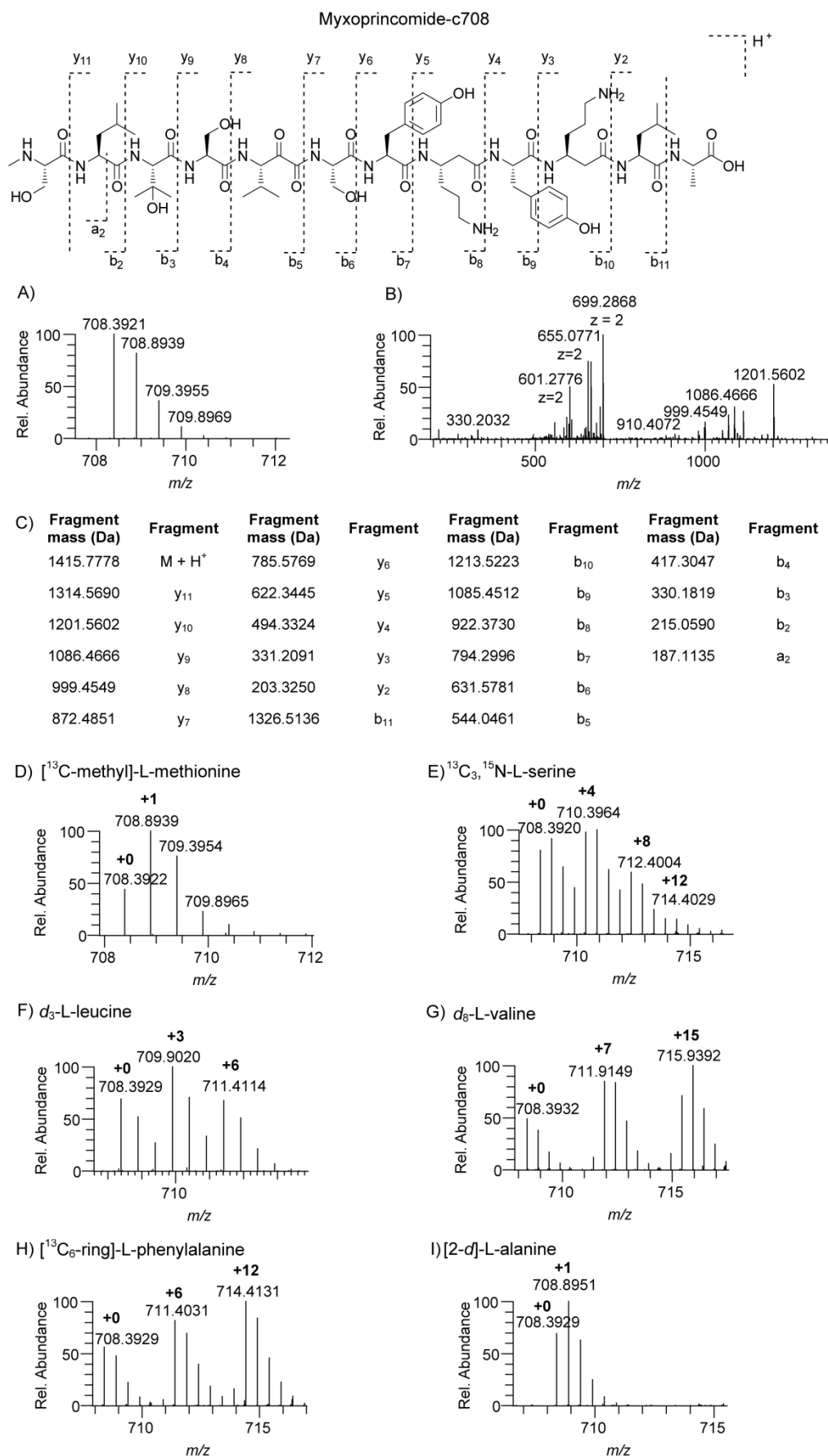
**Figure S4.5:** Myxoprincomide-c651 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds



**Figure S4.6:** Myxoprincomide-c513 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds

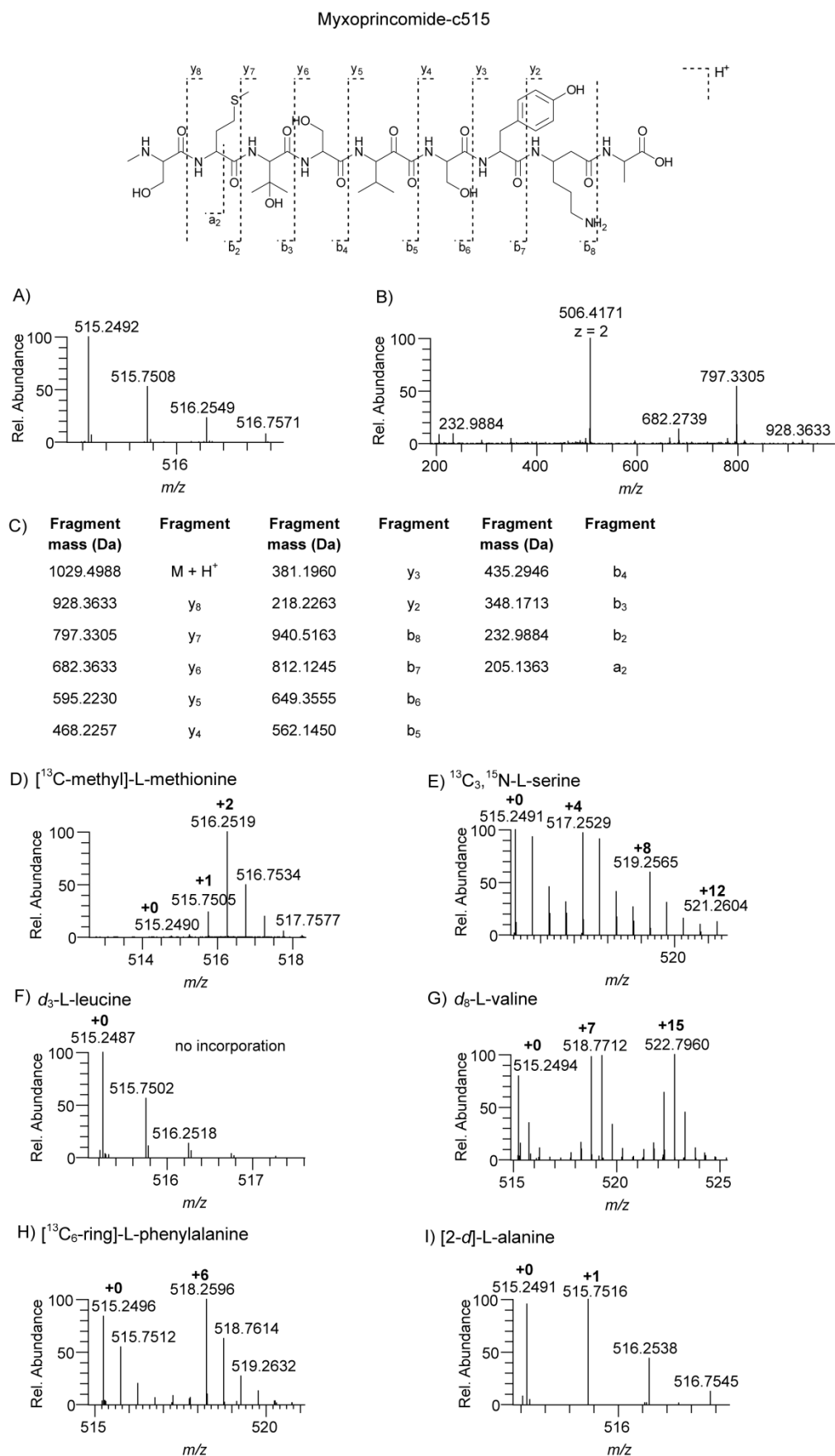


**Figure S4.7:** Myxoprincomide-c562 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds

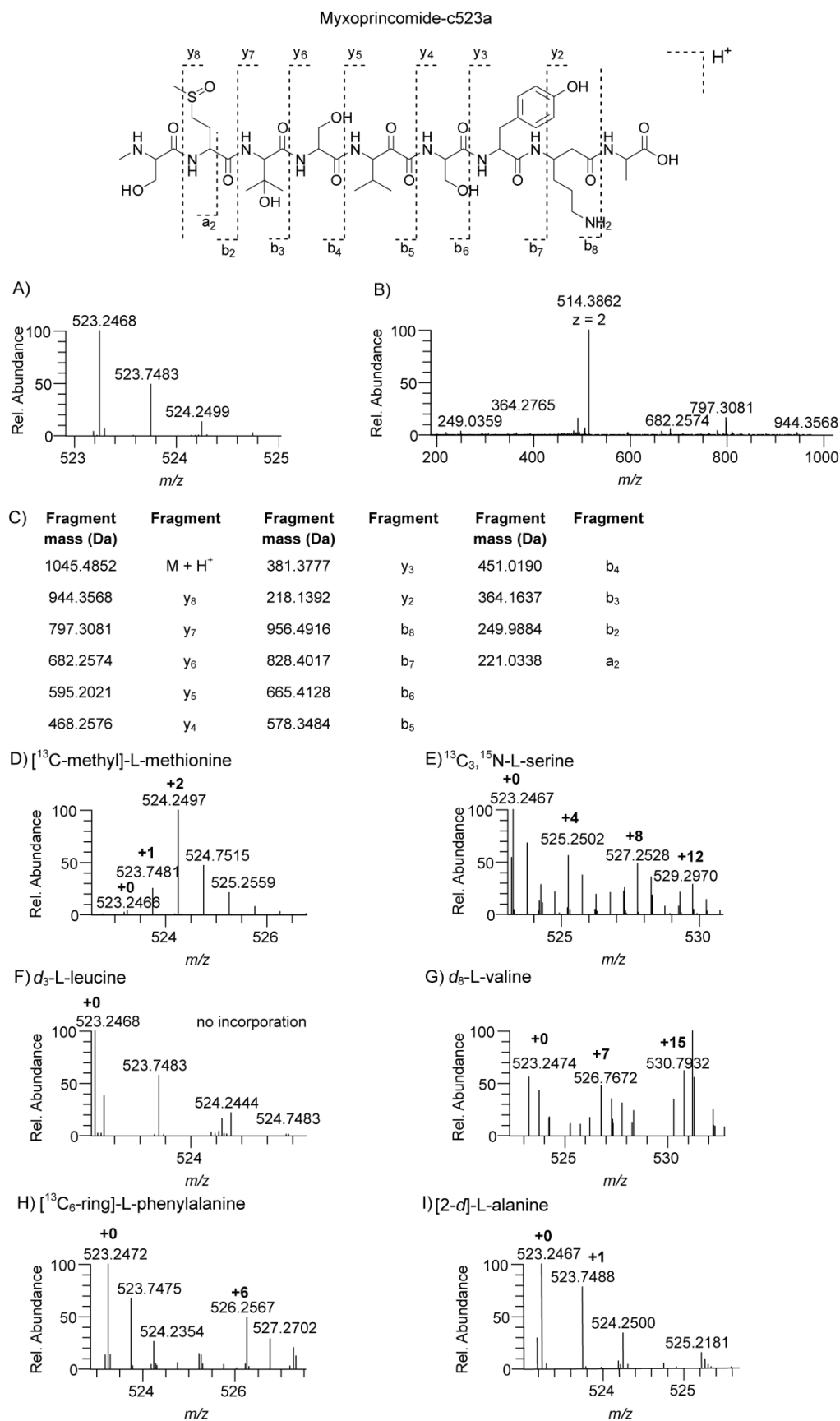


**Figure S4.8:** Myxoprincomide-c708 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds

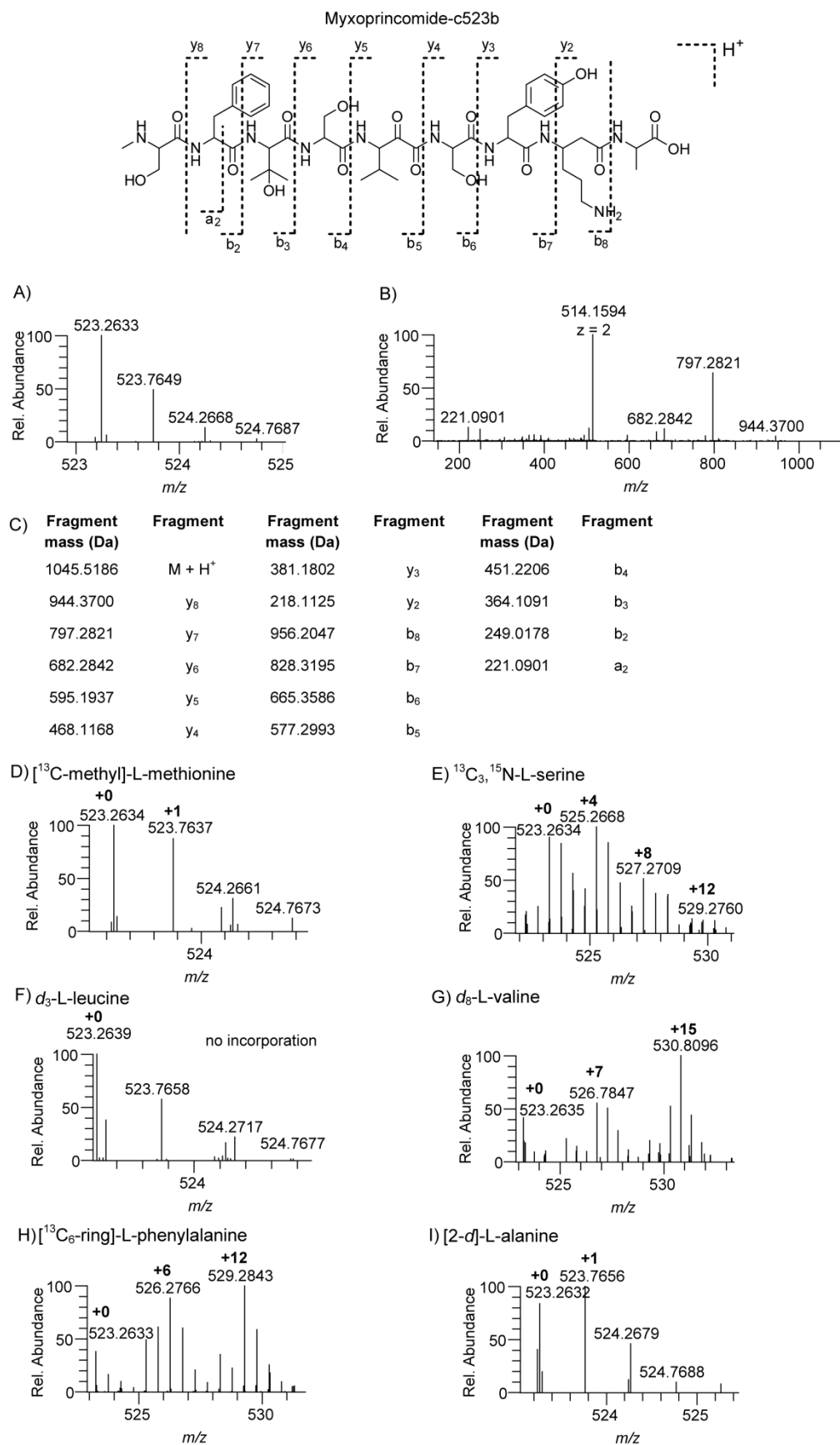




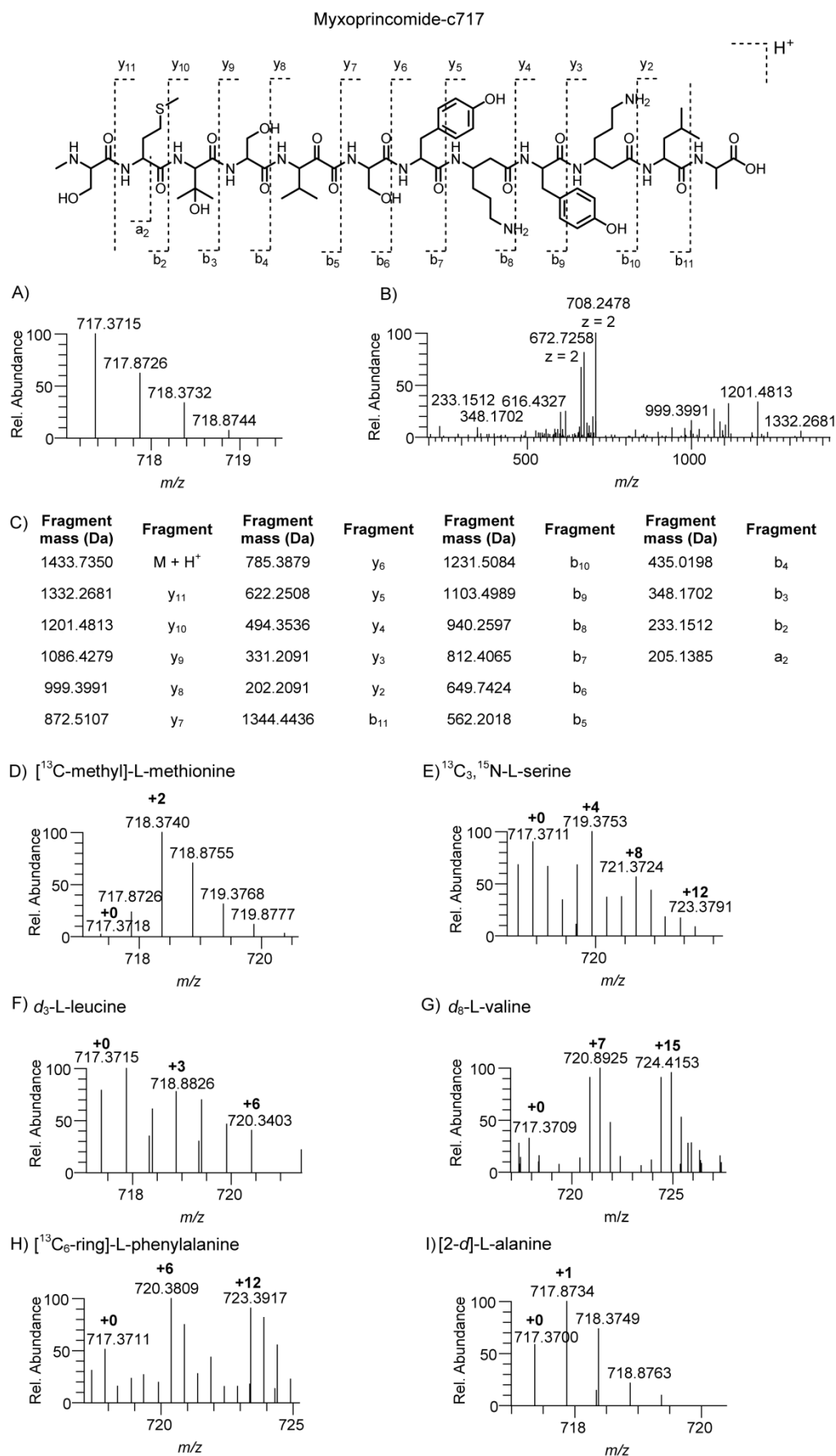
**Figure S4.9:** Myxoprincomide-c515 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds



**Figure S4.10:** Myxoprincomide-c523a MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds



**Figure S4.11:** Myxoprincomide-c523b MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds



**Figure S4.12:** Myxoprincomide-c717 MS data. A) HR-ESI MS spectrum. B) MS<sup>2</sup> analysis. C) Annotated product ions. D-G) HR-ESI MS spectra of labeled compounds

**Table S4.1:** Result of  $^{13}\text{C}$ -acetate labeling of myxoprincomides. Shown are the relative intensities of the isotope peaks in the control (unlabeled),  $1\text{-}^{13}\text{C}$ -acetate,  $2\text{-}^{13}\text{C}$ -acetate, and  $1,2\text{-}^{13}\text{C}$ -acetate labeled myxoprincomides.

Myxoprincomide-c534		Relative Intensity (in %)		
Sample / labeled acetate	534.3124 <i>m/z</i>	535.3160 <i>m/z</i>	536.3188 <i>m/z</i>	
Control	100.00	25.35	4.54	
1- <sup>13</sup> C-acetate	100.00	26.02	4.77	
2- <sup>13</sup> C-acetate	100.00	27.95	5.21	
1,2- <sup>13</sup> C <sub>2</sub> -acetate	100.00	26.03	4.74	
Myxoprincomide-c649		Relative Intensity (in %)		
Sample / labeled acetate	649.3396 <i>m/z</i>	650.3433 <i>m/z</i>	651.3461 <i>m/z</i>	
Control	100.00	30.18	6.25	
1- <sup>13</sup> C-acetate	100.00	29.69	6.40	
2- <sup>13</sup> C-acetate	100.00	40.95	10.73	
1,2- <sup>13</sup> C <sub>2</sub> -acetate	100.00	32.20	7.50	
Myxoprincomide-c651		Relative Intensity (in %)		
Sample / labeled acetate	651.3555 <i>m/z</i>	652.3590 <i>m/z</i>	653.3619 <i>m/z</i>	
Control	100.00	31.14	6.73	
1- <sup>13</sup> C-acetate	100.00	30.71	6.71	
2- <sup>13</sup> C-acetate	100.00	38.92	9.73	
1,2- <sup>13</sup> C <sub>2</sub> -acetate	100.00	32.74	7.50	
Myxoprincomide-c506		Relative Intensity (in %)		
Sample / labeled acetate	506.2715 <i>m/z</i>	506.7729 <i>m/z</i>	507.2758 <i>m/z</i>	507.7767 <i>m/z</i>
Control	100.00	53.32	17.20	3.77
1- <sup>13</sup> C-acetate	100.00	55.07	17.75	4.06
2- <sup>13</sup> C-acetate	100.00	98.35	45.84	15.04
1,2- <sup>13</sup> C <sub>2</sub> -acetate	100.00	68.83	27.37	8.08

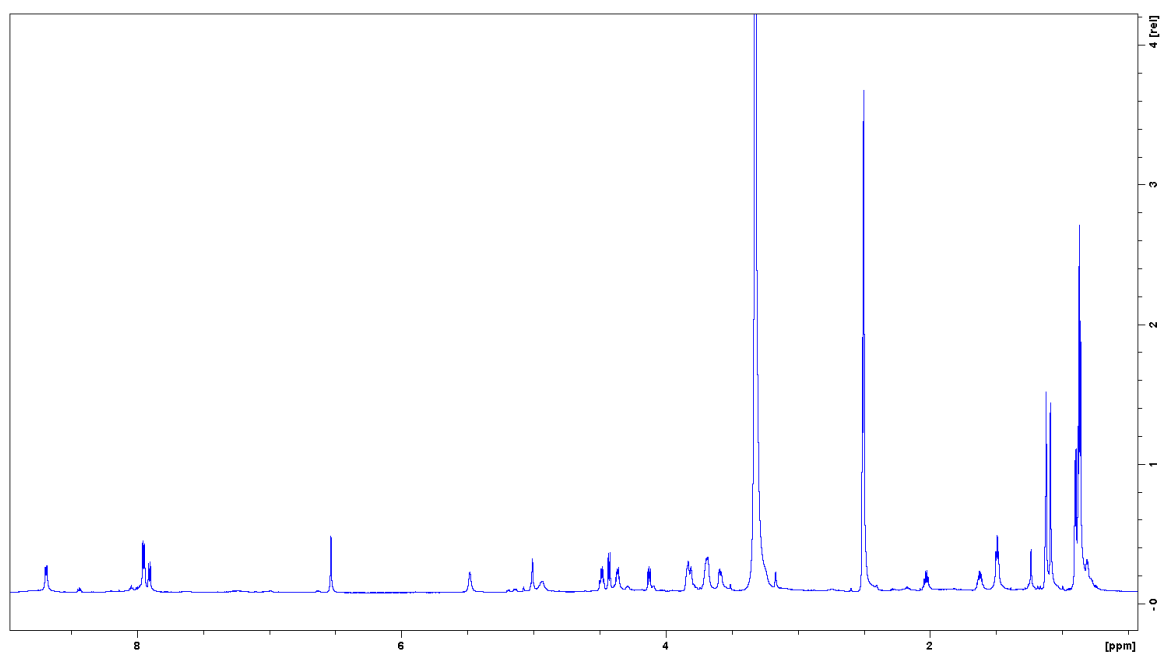
## 4 Myxoprincomides

Table S4.1 (*cont'd*)

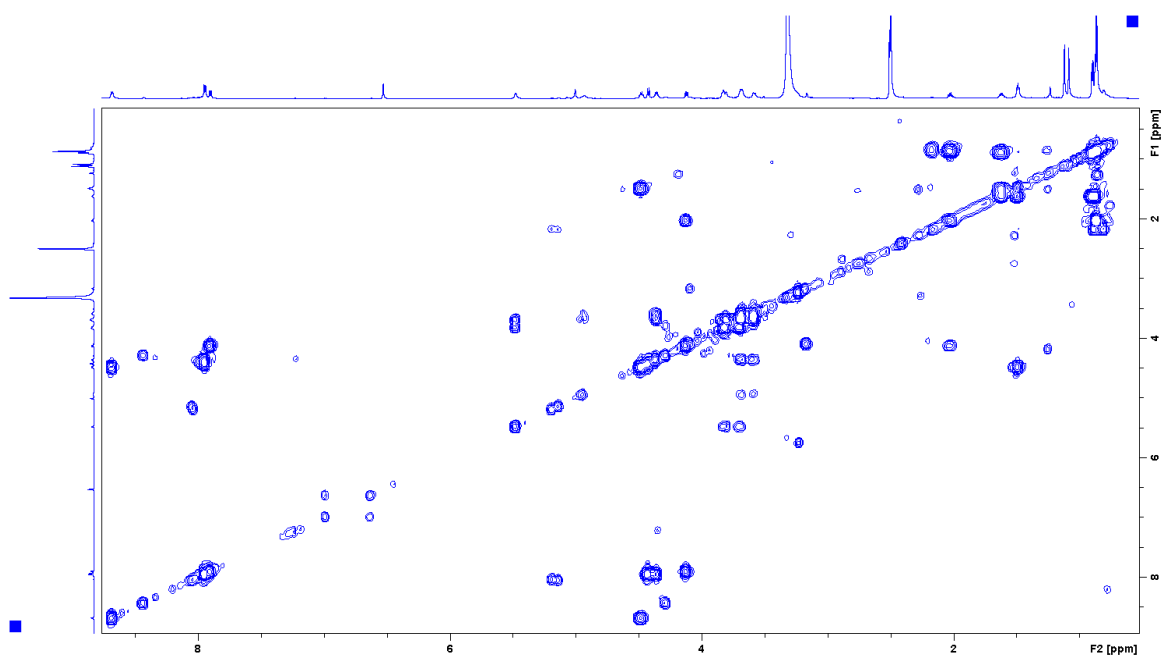
Myxoprincomide-c515			Relative Intensity (in %)			
Sample	/	labeled	515.2759 <i>m/z</i>	515.7775 <i>m/z</i>	516.2789 <i>m/z</i>	516.7804 <i>m/z</i>
acetate						
Control			100.00	51.27	16.69	3.95
1- <sup>13</sup> C-acetate			100.00	54.09	16.32	4.15
2- <sup>13</sup> C-acetate			95.11	100.00	43.76	16.92
1,2- <sup>13</sup> C <sub>2</sub> -acetate			100.00	67.73	27.40	8.01
Myxoprincomide-c523a			Relative Intensity (in %)			
Sample	/	labeled	523.2468 <i>m/z</i>	523.7483 <i>m/z</i>	524.2499 <i>m/z</i>	524.7483 <i>m/z</i>
acetate						
Control			100.00	48.58	12.97	2.87
1- <sup>13</sup> C-acetate			100.00	52.59	16.85	3.84
2- <sup>13</sup> C-acetate			100.00	82.97	32.50	10.81
1,2- <sup>13</sup> C <sub>2</sub> -acetate			100.00	62.08	22.99	3.62
Myxoprincomide-c523b			Relative Intensity (in %)			
Sample	/	labeled	523.2633 <i>m/z</i>	523.7649 <i>m/z</i>	523.2668 <i>m/z</i>	523.7648 <i>m/z</i>
acetate						
Control			100.00	53.71	16.53	2.80
1- <sup>13</sup> C-acetate			100.00	56.82	17.90	4.26
2- <sup>13</sup> C-acetate			99.24	100.00	41.08	12.62
1,2- <sup>13</sup> C <sub>2</sub> -acetate			100.00	78.14	31.29	9.83
Myxoprincomide-c562			Relative Intensity (in %)			
Sample	/	labeled	562.8134 <i>m/z</i>	563.3151 <i>m/z</i>	563.8178 <i>m/z</i>	564.3188 <i>m/z</i>
acetate						
Control			100.00	59.43	25.07	6.47
1- <sup>13</sup> C-acetate			100.00	62.24	21.93	7.68
2- <sup>13</sup> C-acetate			90.47	100.00	48.84	20.10
1,2- <sup>13</sup> C <sub>2</sub> -acetate			100.00	73.03	30.00	11.68
Myxoprincomide-c708			Relative Intensity (in %)			
Sample / labeled acetate			708.3919 <i>m/z</i>	708.8937 <i>m/z</i>	709.3954 <i>m/z</i>	709.8969 <i>m/z</i>
Control			100.00	79.68	37.38	11.13
1- <sup>13</sup> C-acetate			100.00	80.25	37.38	13.22
2- <sup>13</sup> C-acetate			77.16	100.00	65.30	51.20
1,2- <sup>13</sup> C <sub>2</sub> -acetate			100.00	95.81	51.20	20.72
Myxoprincomide-c717			Relative Intensity (in %)			
Sample / labeled acetate			717.5715 <i>m/z</i>	717.3726 <i>m/z</i>	718.3722 <i>m/z</i>	718.8744 <i>m/z</i>
Control			100.00	71.77	32.30	13.96
1- <sup>13</sup> C-acetate			100.00	69.86	38.51	17.59
2- <sup>13</sup> C-acetate			92.22	100.00	62.72	28.63
1,2- <sup>13</sup> C <sub>2</sub> -acetate			97.21	100.00	54.21	22.23

MXAN 4698	-----MLPAGLPPEPTAARASLSAEGRKRLFNPATDAEWADWRWQQRHAVRGL	48
MXAN 4699	MQTTPIRAKPEAGPAEQFFTYPHRREFIEPDWRRIPGFKDVSADWENS VWQRKHTIKNL	60
MXAN 6263	-----MSVMNSAALALSTGRRYRAYTTRHDELTTTRAGLSADERLAV	42
KamA	-----MINR-----RYELFKDVSDADWNDWRWQVRNRIETV	31
SttO	-----MPSW-----ATCSFTRA	11
VioP	-----MEPDWRRLLPGWHDVTAAQWHDVQWQRAHCVKNA	33
	:	
MXAN 4698	EQLERYVP----LTSNERAGVQETSALFRIGISPYLSLIDP-----EHPFC	91
MXAN 4699	KELKATLGALLPDDLAE SMERDQRRERATMSLLVPPQMLNTMN-----LEDLWRD	109
MXAN 6263	QAVAHVLPFRNTSYVDELIDWAAAPADPIYRLVFPQADMLPTEDVARMVD--LLSSGAS	100
KamA	EELKKYIP----LTKEEEEGVACQVKSLRMAITPYLSLIDP-----NDPND	74
SttO	-----FYEDWEWDRSHYATMSVLLPPQMINTMAPEVAAARPGLTKAFYDD	57
VioP	RQLRAVVGDLDDKFYDDLTDQE HMATMAMLITPQMLNTIAPETPADSDG-YHDAFYAD	92
	: . : . :	
MXAN 4698	PVRMQSIPVRAEAR----IRPGE LADPLGEDKTRPEECIVHKYPDRVLFALD--TCSVY	145
MXAN 4699	PVRRYMLPAYADRLTEWTNHPKASRDSLHEQDMWVVEGLTHRYPTKVLAEMLP--TCPQY	167
MXAN 6263	PLELNAAANEIRARLNPHFAGQMQLNVPKLANE EFPVGLQHKYKETVLIFPKQQTCHAY	160
KamA	PVRKQAIPTALELN---KAAADLEDPLHEDTDSVPVGLTHRYPDRVLLITD--MCSMY	128
SttO	PVRRYMPVMSDRHPQWPSHPMASRDSLHEQDMWVVEGLTHRYPTKVLAELLS--TCPQY	115
VioP	PVRRYMPVRSRDLRWPSHPLSSRDSLHEAEMWVVEGLTRRYPTKVLAELVA--TCPQY	150
	*:. . : : : * * *	
MXAN 4698	CRHCTR--RRITQGGVAELSK-----EQLRRGVDYVRSHPEVRDVLISGGDPFMLS DS	196
MXAN 4699	CGHCTR--MDLVGNDVPQVSKHKFGIGPKDRYAQMLDYLRRTPTRD VVVSGGDIANLP IQ	226
MXAN 6263	CTYCFRWAQFVGADLKFAS-----REIEPLVNYIRAHPEVTNVLFTGGDP MIMTEA	212
KamA	CRHCTR--RRFAGQSDDSMPM-----ERIDKAIDYIRNTPQVRDVLISGGDALLVSDE	179
SttO	CGHCTR--MDLVGNSTPQVTKSRLQLKPVDRA DRILAHLRVSPGIRDVVVSGGDLANMPWP	174
VioP	CGHCTR--MDLVGGSTPSVDKQRLTLRPADRQEAILDHLRRTPGVRD VVVSGGDVANVPWP	209
	* : * * : . . : : : * * : : : * * : : : * * :	
MXAN 4698	RLEELLAPLSEIPHVEMIRIGTRVPVCLPMRVTD-----ALAKTLRRYAPVFV VTH	247
MXAN 4699	QLEPFVSSLM DIPNIRDIRLASKGLMAIPQHFLQDSVLQGLDRLAKKAVERGV D LALHTH	286
MXAN 6263	VLAKYIEPLLDIEHLEAIRIGTKALAYWPRFVTDS DADDILRLF EKVVASGKSLAFMAH	272
KamA	TLEYIIAKLREIPHVEIVRIGSRTPVVL PQRITP-----ELVNMLKKYHPVWLNTH	230
SttO	RLERFLDDLLEIDSIRDIRLASKALIGLPQHWN SGP LLEGVARIARKARSRVRIALHTH	234
VioP	RLESFLLRLEIDSVRDIRLASKALVGLPQH WLPQVVSGLENVAGVAARRGVHLAVHTH	269
	* : * * : . . : : : * : : : : *	
MXAN 4698	FNHPKEVTP EAREACERLVDHGVP--VENQAVLMRQLNSDARI IKELSHLLLRSR--VRPYY	305
MXAN 4699	VNHAQQLTPLVGKAVRKLL EMGFRDVRNQGVLLRGVNDSPQALLDLCFTLLDHAKILPYY	346
MXAN 6263	FSHPNEMVPEIVQEAVRRIRGTGAVIRTQAPLIR TINDTPGTWESMWRTHLRHG--MVPYY	331
KamA	FNHPNEITEESTRACQLLADAGVP--LGNQSVLLRGVND CVHVMKELVNKLVKIR--VRPYY	288
SttO	ANAAQQVTPAVARA AAWALLDAGLHDVRNQGVLMRGVND SAHDL D LCFALTDHAGITPYY	294
VioP	ANHVSQSVTPLVAEGARALLDAGVRDVRNQGVLMRGVND STAALLDLCFALQDEAGILPYY	329
	. : : . . : : * . * * : * . : : : * * :	
MXAN 4698	LHQM DVAEGCEHLRTPIAKGLEIIQQLRGYTTGLAVPHLAVDLPGGGGKVTLPQPDYAVEY	365
MXAN 4699	FYMC DMIPNSEHWRLSVAQAQQLQHDIMGYMPGFATPRIVCDVPFVG---KRWVHQVAEY	403
MXAN 6263	MFVERDTGPQDYFAVPLAEAYDIFRNAFQSVSGLARTVRGSPMSATPG---KVCVDGVAEI	389
KamA	IYQCDLSLGLEHFRTPVSKGIEIIEGLRGHTSGYCVPTFVV DAPGGGGKTPVMPNYVISQ	348
SttO	FYMC DMIPNAEHWRVPLHRAQLIQRQIMGYLPGFATPRIVCDVP MAG---KRWVDQDTSY	351
VioP	FYMC DMVPGA EHWR TSLAEAQDLQHAIMGYLPGYATPRIVCDVPYVG---KRWVHQVAEY	386
	: . : : : . . : . . * . . . . . :	
MXAN 4698	GAQETVFRNYKGERFTYPEPEETDCTCPYDG--VWQERAPRYGYR-----KG-----	410
MXAN 4699	DRERGISYWKNYRTGIELNDADALS RKEYEYFDPIDQLSEAGQAWWR---EQQKAA----	456
MXAN 6263	AGEKVFLHF IQSRDPELVGRPF FAKYDEKASWLFDLKPAMGATHFPWEEPSA-----	442
KamA	SHDKVILRNFE GVIITYSEPINYPGNCNDVCTGKKKVHKVGVAGLLNNEGMALEPVGLE	408
SttO	DRELGVSHWSKSYLTPL EAADRDAHDGTYHYDPI DTLPLD GQRWWR---DIRQG-----	403
VioP	DRERGISYWKNYRTAIELDDPDALTRYPFHDPLSTLPESGWRWWRQVAAREGQACA-	445
	: . *	
MXAN 4698	-----	
MXAN 4699	-----	
MXAN 6263	-----	
KamA	RNKRHVQE	416
SttO	-----	
VioP	-----	

**Figure S4.13:** Alignment of MXAN\_4698, MXAN\_4699, and MXAN\_6263 and known 2,3-lysine amino mutases KamA (AAD43134), SttO (CAB67710), and VioP (AAP92506.1). Highlighted are the conserved motifs for 2,3-LAM activity: CxxxCxxx (yellow), Arg<sub>134</sub>-Arg<sub>293</sub>-Asp<sub>300</sub> (turquoise), and Lys<sub>337</sub> (purple).



**Figure S4.14:**  $^1\text{H}$  NMR spectrum of myxoprincomide-c534 in  $\text{DMSO}-d_6$



**Figure S4.15:** COSY NMR spectrum of myxoprincomide-c534 in  $\text{DMSO}-d_6$



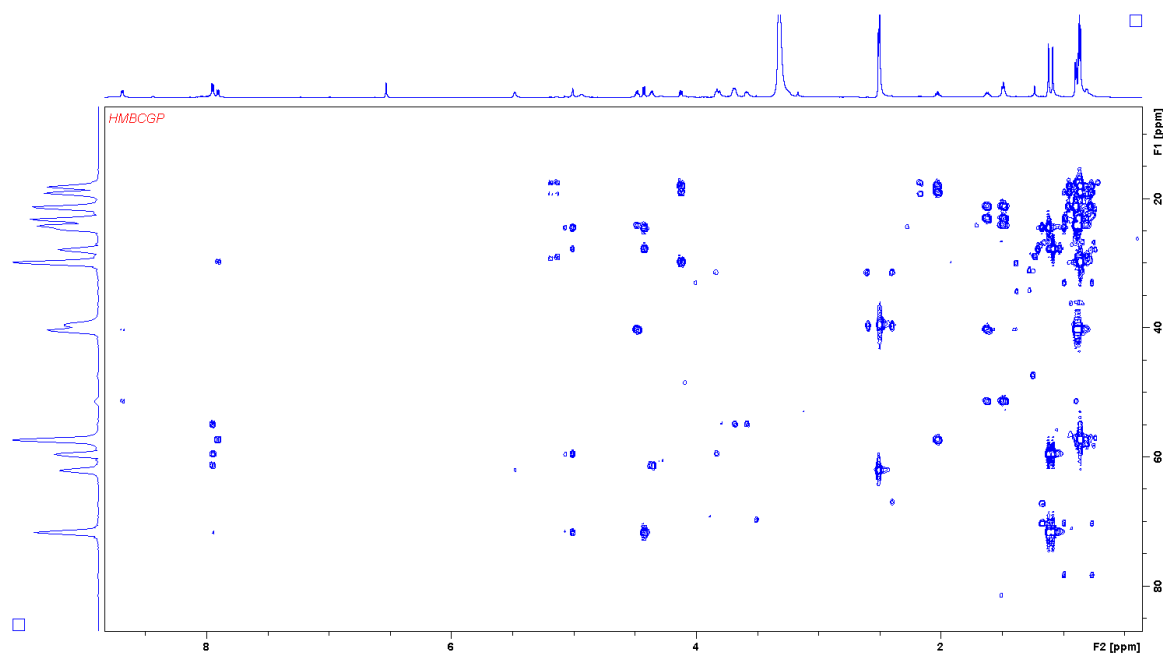


Figure S4.16: HMBC NMR spectrum of myxoprincomide-c534 in DMSO- $d_6$

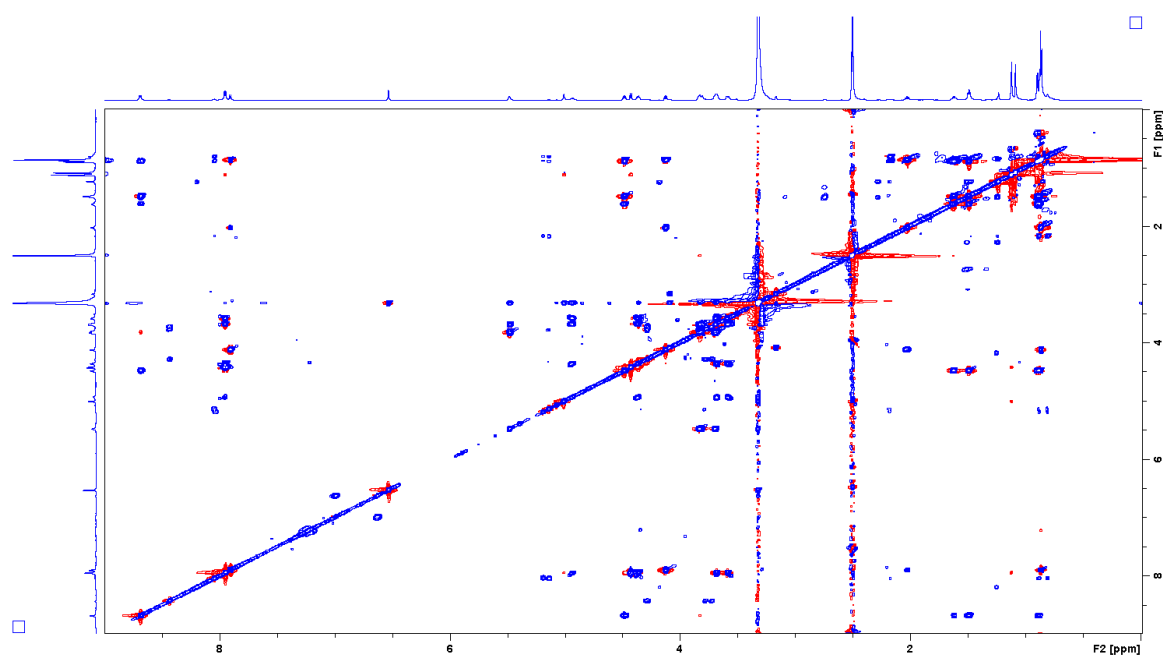
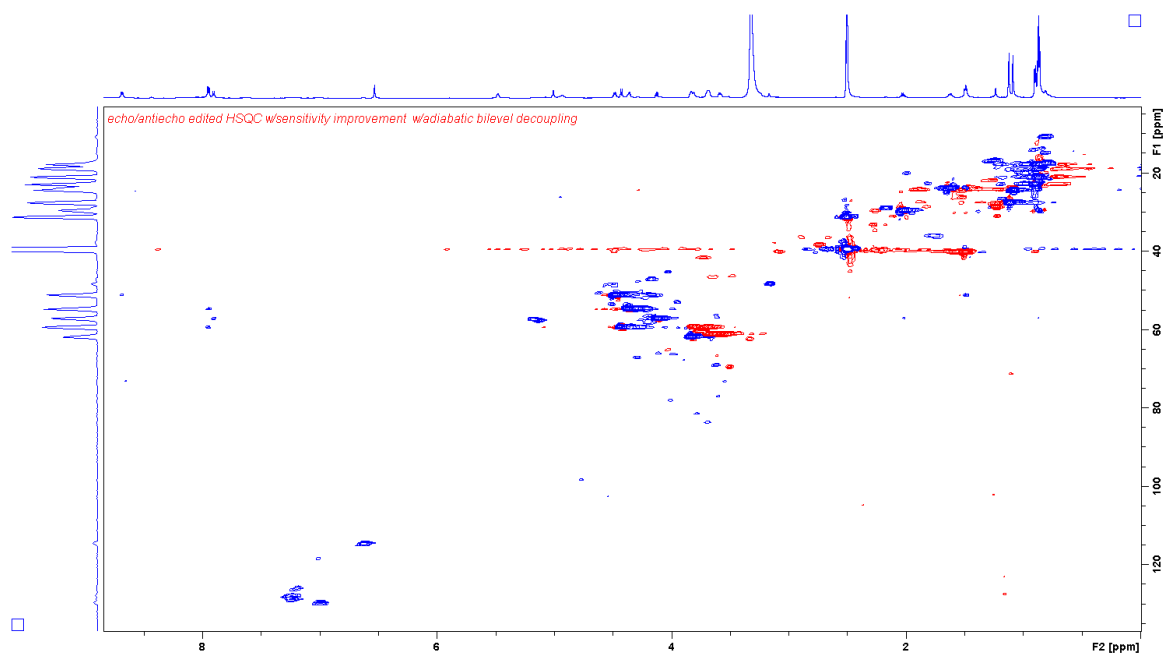
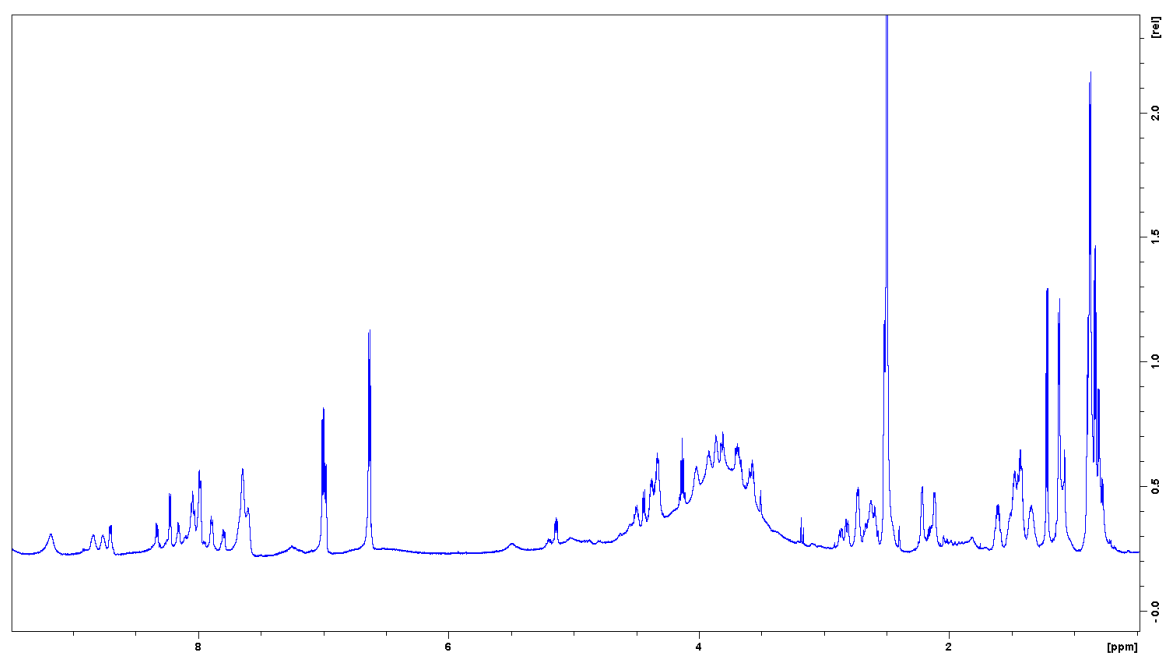


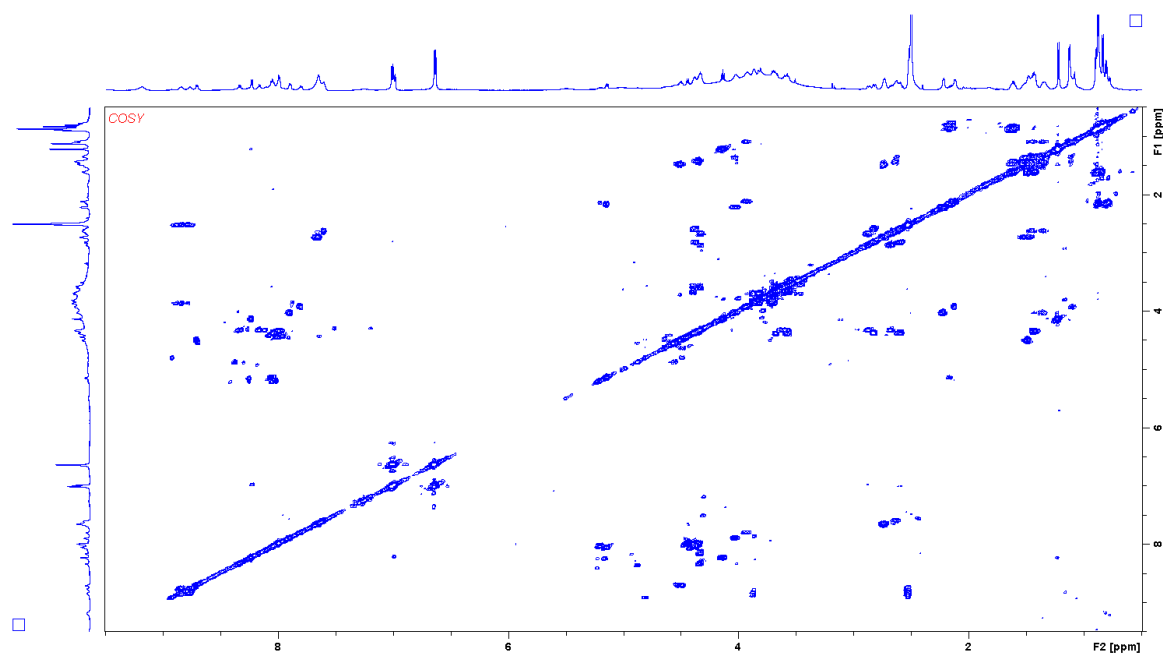
Figure S4.17: HOHAHA spectrum of myxoprincomide-c534 in DMSO- $d_6$



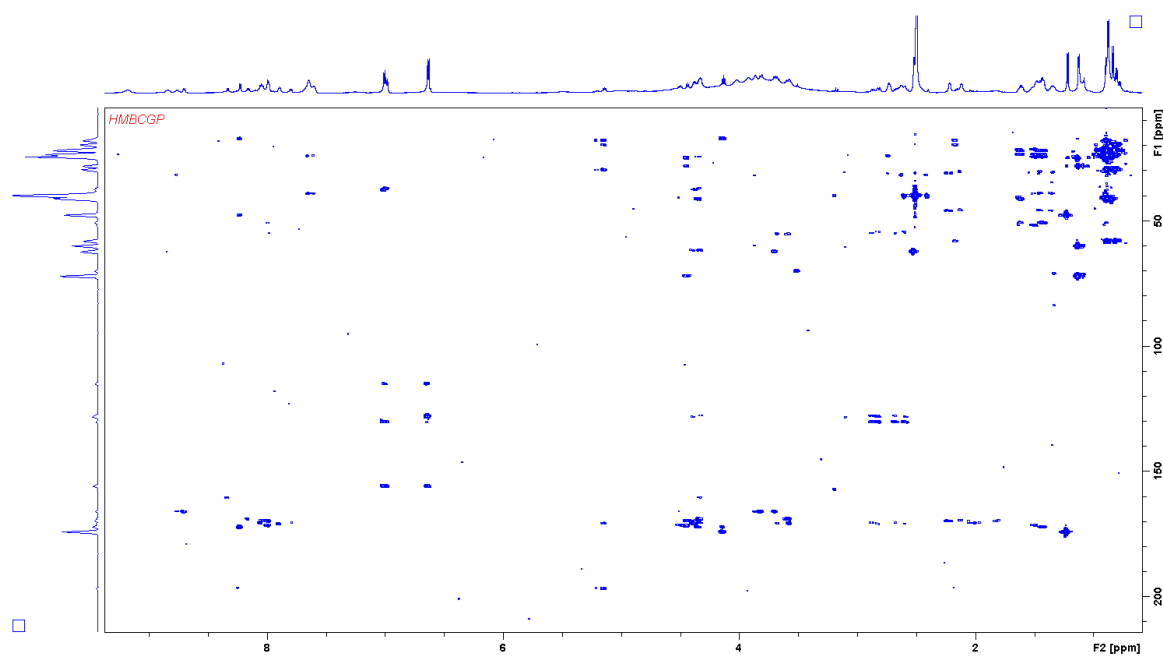
**Figure S4.18:** HSQC NMR spectrum of myxoprincomide-c534 in DMSO- $d_6$



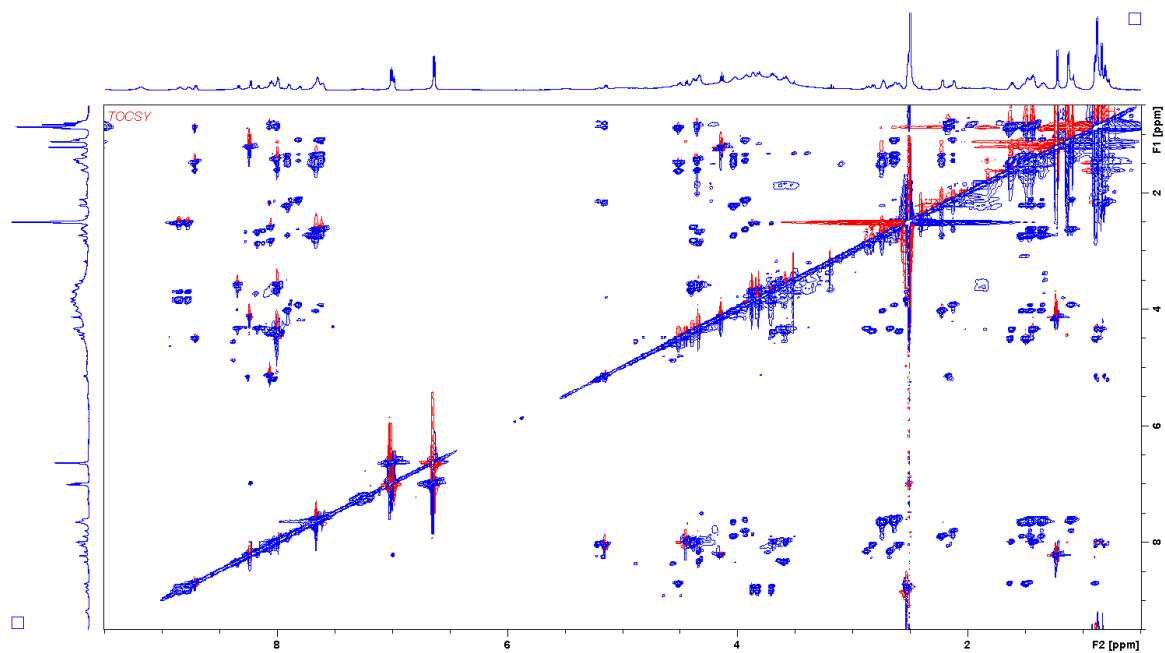
**Figure S4.19:**  $^1\text{H}$  NMR spectrum of myxoprincomide-c708 in DMSO- $d_6$



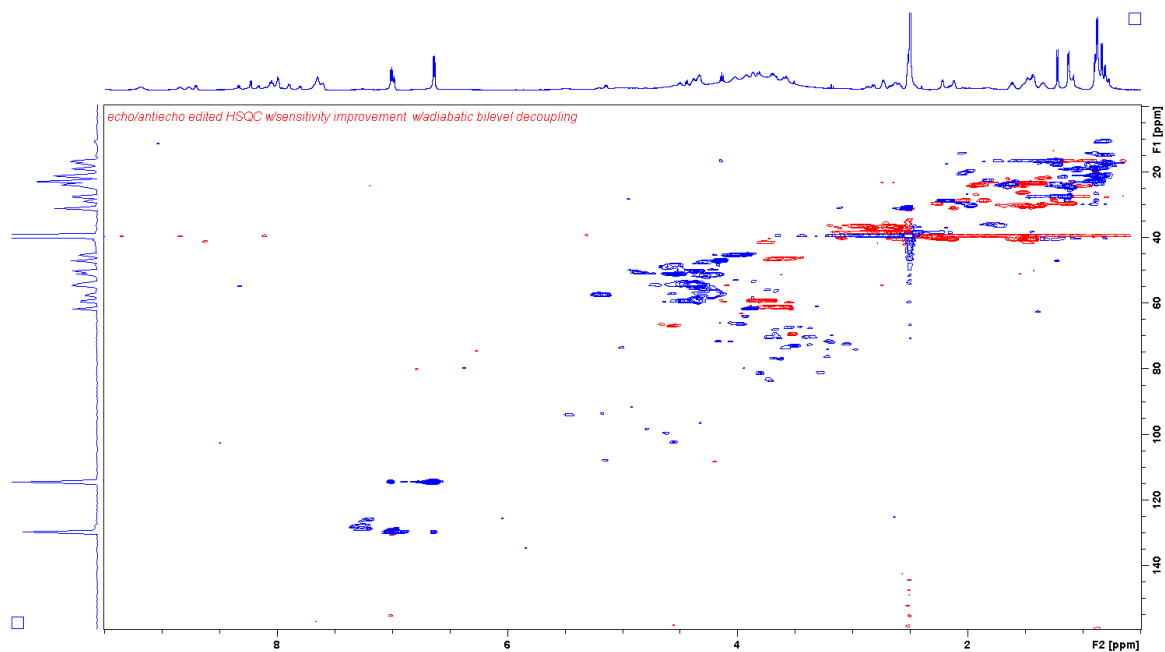
**Figure S4.20:** COSY NMR spectrum of myxoprincomide-c708 in DMSO- $d_6$



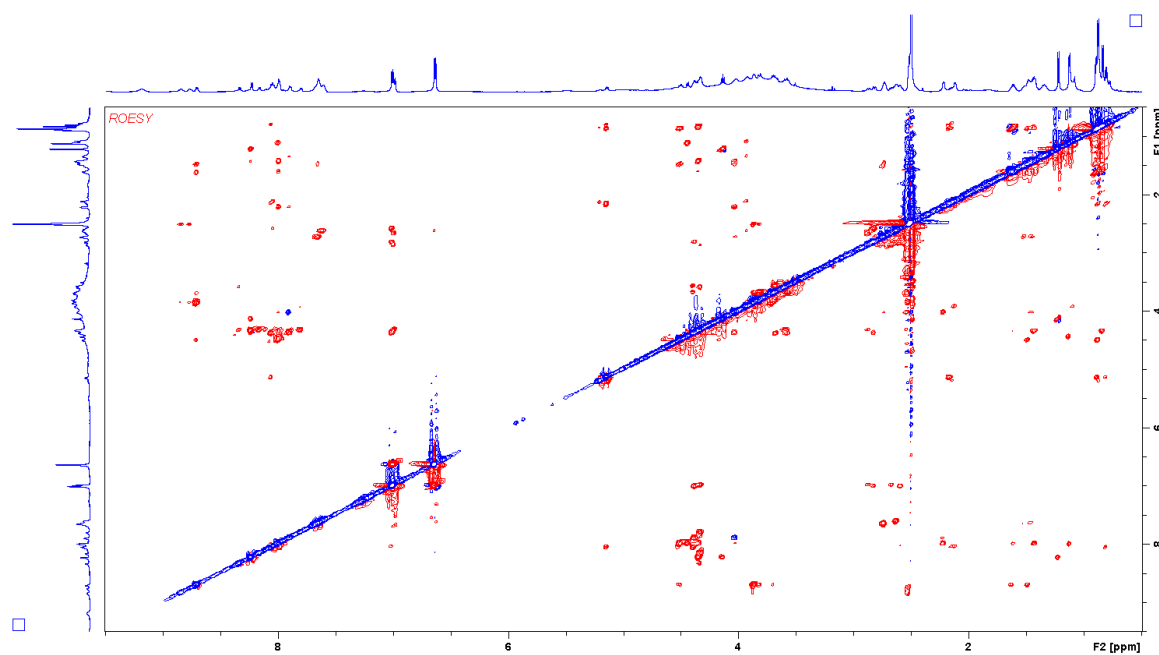
**Figure S4.21:** HMBC NMR spectrum of myxoprincomide-c708 in DMSO- $d_6$



**Figure S4.22:** HOHAHA NMR spectrum of myxoprincomide-c708 in DMSO- $d_6$

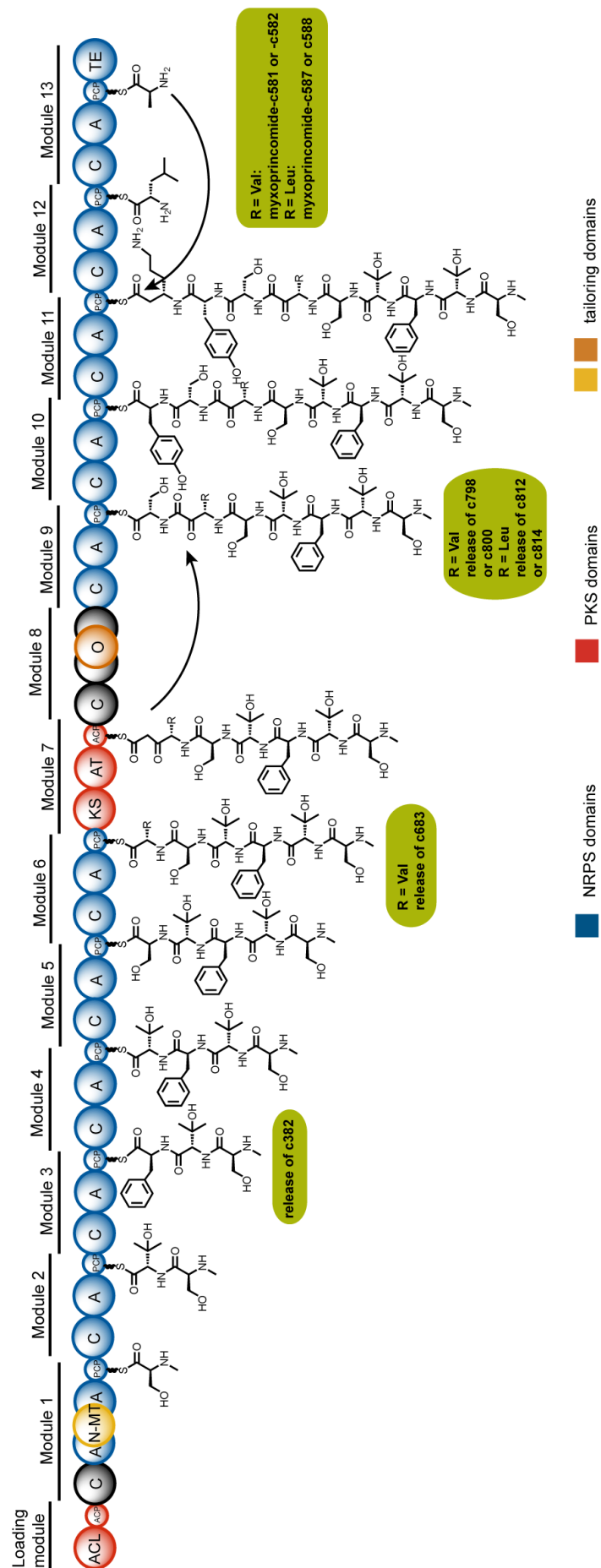


**Figure S4.23:** HSQC NMR spectrum of myxoprincomide-c708 in DMSO- $d_6$



**Figure S4.24:** ROESY NMR spectrum of myxoprincomide-c708 in DMSO- $d_6$

## 4 Myxoprincomides



## 5 Identification and characterization of the althiomycin biosynthetic gene cluster in *Myxococcus xanthus* DK897

Cortina, N. S., Revermann, O., Krug, D., Müller, R. *ChemBioChem* **2011**, 12, 1411 – 1416.

This article is available online at:

<http://onlinelibrary.wiley.com/doi/10.1002/cbic.201100154/pdf>

The supporting information is available online at:

[http://onlinelibrary.wiley.com/store/10.1002/cbic.201100154/asset/supinfo/cbic\\_201100154\\_sm\\_miscellaneous\\_information.pdf?v=1&s=ffe6740f9ea509a8a741bb33bca9b6f0fc5658af](http://onlinelibrary.wiley.com/store/10.1002/cbic.201100154/asset/supinfo/cbic_201100154_sm_miscellaneous_information.pdf?v=1&s=ffe6740f9ea509a8a741bb33bca9b6f0fc5658af)

## 6 Conclusion & Outlook

The main achievement of this work is the association of four secondary metabolites produced by *Myxococcus xanthus* to their corresponding biosynthetic pathways. These compounds include three novel compound classes from *M. xanthus* DK1622, including the myxoprincomides and the known broad spectrum antibiotic althiomycin as produced by *M. xanthus* DK897. A full *de novo* structural elucidation for the myxoprincomide compound family is reported. Based on these findings, models for althiomycin and myxoprincomide biosynthesis by *M. xanthus* are presented.

### 6.1 Secondary metabolome mining

The effectivity of coupling gene-inactivation by targeted mutagenesis with comprehensive analysis of HPLC-high resolution MS data sets to unearth the products of orphan pathways was clearly shown in this work. Three unassigned biosynthetic pathways from *M. xanthus* DK1622 were correlated to their associated products via a straightforward methodology, i.e. targeted gene inactivation, one cultivation condition, one extraction method, one separation condition, and one MS ionization mode. It is an extended version of the targeted mutagenesis-metabolite profiling method in that instead of manual inspection of HPLC-HR MS data sets, a comprehensive analysis of the extract profile is performed by principal component analysis following an advanced compound finding algorithm. The development and application of the compound finding algorithm and PCA to uncover the “hidden” secondary metabolome of myxobacterial strains was previously done and allowed to define a panel of hypothetical compounds by high-resolution mass values and calculated putative molecular formulae.<sup>[1, 2]</sup> However, in the present study the underlying “comprehensive metabolomics” analytical platform was applied to directly correlate select biosynthetic pathways to their corresponding products.

Using this approach, three new secondary metabolites—c506, c844, and c329—were identified and correlated to their biosynthetic gene clusters in *Myxococcus xanthus* DK1622 (c506 : MXAN\_3779,  $[M+2H]^{2+}$  506.27  $m/z$ ), c844 : MXAN\_1599 – MXAN\_1608,  $[M+H]^+$  844.37  $m/z$ ) and c329 : MXAN\_3628 – MXAN\_3633,  $[M+H]^+$  = 329.19  $m/z$ ).



Amongst the three compounds, further work on compound c506, termed myxoprincomide-c506, was carried out. Myxoprincomide-c506 was an interesting compound to investigate as its biosynthetic pathway MXAN\_3779 was the largest, continuous hybrid synthetase from DK1622 capable of assembling an entire structural backbone.

Using semiquantitative *M. xanthus* species-wide metabolomics data, *M. xanthus* A2 was determined to be a better producer of myxoprincomide-c506. In addition to optimizing cultivation conditions to generate a higher yield, *M. xanthus* strain A2 was genetically engineered to artificially overexpress its MXAN\_3779 homologue by insertion of the T7A1 promoter upstream of the gene resulting in the optimized strain A2.c506 which can produce myxoprincomide-c506 at levels 30 times higher than DK1622. Large-scale cultivation and fractionation of A2.c506 enabled the isolation of 1.2 mg of pure myxoprincomide-c506 allowing for structural elucidation by extensive 2D NMR spectroscopy and ESI-MS/MS techniques.

After discovery, access to the purified compound is the next challenge to be addressed for a complete secondary metabolite structural characterization. The case of myxoprincomide-c506 is a clear example of the efforts required to obtain sufficient amounts of pure low abundance compounds. On top of cultivation optimization, genetic manipulation by insertion of a strong constitutive promoter contributed to the overall product yield enhancement. Fortuitously, the *M. xanthus* strains used in this work (DK897 and A2) were genetically-tractable and the transformation protocol for *M. xanthus* DK1622 was intraspecifically-suitable. However, this is not necessarily the case for other myxobacterial strains. The need to tailor, establish, and improve molecular biological tools to perform targeted mutagenesis benefits not just the correlation of pathways to their products but also the improvement of secondary metabolite production by genetic manipulation.

Further work is required to isolate and determine the final structures of the DK1622 compounds c844 and c329. The same amount of effort placed into the determination of the final structure of myxoprincomide-c506 should be done for the structure elucidation of c844 and c329. Initial investigations have been done in this regard, particularly in determining the best native producer for each of the compounds (data not shown). Whereas the biosynthetic gene cluster for c329 appears to be encoded in one operon,

which can be easily subjected to yield optimization by insertion of a constitutive promoter, the biosynthetic gene cluster for c844 (MXAN\_1600 – MXAN\_1608) is not and therefore requires regulation of gene expression facilitated by, for example, enhancer binding proteins, to alter its production.<sup>[3]</sup>

Although this work was successful in expanding the secondary metabolite portfolio of *M. xanthus*, further efforts are required to associate the remaining 10 unassigned biosynthetic pathways to their corresponding products. For unidirectional gene clusters, a statistical comparison of the LC-MS profiles between an overexpression vs. knockout mutant is a promising strategy to extricate new compounds. Modification of cultivation conditions by changing the growth medium (e.g. to VY/2) or by exposing the strains to environmental stress (i.e. higher or lower growth temperature or pH) could also impact secondary metabolite production. As opposed to the endpoint-to-endpoint method used in this work, a kinetics approach where strain culture replicates are harvested at different time points during cultivation could provide insight into compounds produced early during bacterial growth but are consumed or degraded at later stages of the growth cycle.

Extraction procedures could be expanded by using other solvents aside from methanol to extract the compounds adsorbed on the XAD-16 resin. Whereas methanol is an excellent solvent, it was noticeable that the XAD-16 resin was still significantly colored post extraction implying that there are still a significant amount of compounds bound to the resin. A methanol / acetone solution could potentially extract more compounds. One caveat is the amount of artifacts normally observed when using acetone as an extraction solvent. One could also conduct a serial extraction using hexane, ethyl acetate, and methanol as three separate extraction stages.

Utilizing a buffer with a higher pH during separation and shifting to negative ionization mode could unearth more compounds which do not ionize well under acidic conditions. For the data already on hand, lowering the stringency of the molecular feature finding parameters could reveal more compounds. This has the risk of increasing the complexity of the bucket table but it does increase the chance of including the low abundance compounds whose mass peak clusters have a lower number of consecutive scans (smaller compound length).

## 6.2 The Myxoprincomides from DK1622

In the process of optimizing the production of myxoprincomide-c506, the structural diversity of the myxoprincomides was revealed. Fifteen additional myxoprincomide variants were identified and categorized into four different groups. Group I-A is composed of myxoprincomide variants which are shorter than the representative structure myxoprincomide-c506. Myxoprincomide-c506 is part of Group I-B which are composed of structures with L-leucine at position 2 of the structure. Group I-C on the other hand are made up of myxoprincomide variants with different residues at position 2 (i.e. methionine, methionine sulfoxide, phenylalanine). Group I-D differentiates itself from the rest by lacking the 417  $m/z$  product ion in its MS<sup>2</sup> spectra.

Out of the 15 additional variants observed, 11 variants were structurally elucidated by combining isotope-labeled substrate feeding, spectrometric, and spectroscopic analyses. Further efforts are required to carry out the structural elucidation of myxoprincomide-c535, -c591, -c555, and c-584 as their low concentration have prevented tandem MS analysis.

The structural depth of the myxoprincomides further established the extent by which Mxp<sub>1622</sub> biosynthetic pathway deviates from strict colinearity. The advantage of using the comprehensive metabolomics approach in discovering this novel myxobacterial compound family is clear in that the actual product spectrum could not have been predicted by genome mining and sequence analysis alone. In addition, the myxoprincomide structural diversity highlights the unusual features of the Mxp<sub>1622</sub> assembly line. Adenylation domain substrate promiscuity, formation of the  $\alpha$ -keto functional group, shunt product release, an irregularly functional module (module 11), and retromodular stuttering are among the biosynthetic steps peculiar to Mxp<sub>1622</sub>.

The revised and extended biosynthetic model provided in Section 4.4.4 is a sound, hypothetical proposal as to how the myxoprincomides are assembled. Nevertheless, empirical evidence by *in vivo* and *in vitro* analysis are necessary to confirm the proposed biosynthetic steps.

The extent of the substrate promiscuity of Mxp<sub>1622</sub> module 2 A domain can readily be determined using the classical ATP-pyrophosphate exchange assay.<sup>[4]</sup> The same assay can establish the kinetic proof as to why L-leucine is the preferred substrate over phenylalanine and methionine.

The  $\alpha$ -keto moiety observed in almost all of the myxoprincomides is a rarely observed functional group in secondary metabolites.<sup>[5-8]</sup> Mxp<sub>1622</sub> is the first reported molecular assembly line to contain an enzymatic module which can form this unusual functional group. Elucidation of the mechanism of the PKS-derived formation of the  $\alpha$ -keto moiety could be achieved by *in vitro* experiments using a recombinant partial assembly line (including Mxp<sub>1622</sub> PKS module 6 and NRPS module 7) and synthetic intermediates. The success of this *in vitro* assay is admittedly dependent on the behavior of the recombinant protein (i.e. proper folding, solubility) and determination of the proper assay conditions. Alternatively, *in vivo* experiments using the optimized producer strain A2.c506 and feeding with <sup>18</sup>O<sub>2</sub> during fermentation could provide insight as to the source of oxygen and the possible mechanism by which the oxidation occurs.<sup>[9]</sup> Also insertion of a promoter upstream of modules 6 and 7 separately to form a truncated assembly line and feeding with activated synthetic intermediates could provide further understanding about the assembly, assuming that the truncated assembly line is still catalytically functional. The role of the Ox domain in the installation of the  $\alpha$ -keto functional group could also be determined by a point mutation or an alanine scan of the Ox domain, providing mutants with a defective Ox domain which could generate myxoprincomide variants with the intact acyl extension in the middle of the peptide.

The group I-A myxoprincomides were shown to be legitimate shunt products when these compounds were detected in the crude extract after the gene encoding the 2,3-lysine aminomutase (MXAN\_4699), which provides the  $\beta$ -lysine for module 10 of Mxp<sub>1622</sub>, was inactivated. Fusion of the TE domain upstream of the later modules forming a truncated Mxp<sub>1622</sub> could provide further insight into the release mechanism of the shunt products. This method has been shown to release intermediates from polyketide synthase assembly lines. <sup>[10, 11]</sup>

The discovery of the larger myxoprincomides (-c562, c708, and c717) which contain an L-leucine residue in the penultimate position in relation to the C-terminus conceded that module 11 is still a functional module. This revelation is contrary to the lack of L-leucine prior to L-alanine in majority of the observed myxoprincomides. It is possible that module 11 of the Mxp<sub>1622</sub> pathway is a sluggish module which impart three biosynthetic scenarios: 1) Module 11 is skipped to form most of the myxoprincomide structures. 2) module 11 works and installs L-leucine to create myxoprincomide-c562. 3) Module 11 is stalled and instead of skipping, the chain extension reverts back to module 9 in a

biosynthetic step termed here retromodular stuttering and proceeds with a L-tyrosine, L- $\beta$ -lysine, L-Leucine, and L-alanine extension to form myxoprincomide-c708 and -c717. Isomodular stuttering has been previously reported for NRPS pathways (e.g. SfmC of the saframycin pathway is used twice to install a second 3-hydroxy-5-methyl-*O*-methyl-tyrosine to complete the tetrapeptide backbone and Cgc18 A domain is used iteratively to load both Cgc18\_PCP in *cis* and Cgc19\_PCP in *trans*)<sup>[12, 13]</sup> However, retromodular stuttering in Mxp<sub>1622</sub> during chain elongation is a rare occurrence in NRP biosynthesis and it is worth to explore the mechanism by which this remarkable phenomenon takes place.

The release of shunt products and the occurrence of retromodular stuttering raise the question of the catalytic efficiency of Mxp<sub>1622</sub>. *In vivo* translation of large multienzyme complexes and the consequent secondary metabolite biosynthesis are metabolically costly for the microorganism. Therefore, if there are any rate limiting steps that lower the catalytic efficiency of Mxp<sub>1622</sub>, it is possible that the pathway compensates by releasing the intermediates, skipping modules, or reusing the upstream modules. The formation of the  $\alpha$ -keto moiety possibly stalls the nascent chain at module 6. (Fig. 4.25) Instead of proceeding with the PKS extension and  $\alpha$ -keto formation, and with the pressure of the next round of peptide biosynthesis pushing the assembly, myxoprincomide-c543 is released. For the growing peptide which actually proceeds through modules 6 to 8, the nascent chain could once again be stalled at module 8 while module 9 is reused for retromodular stuttering resulting in the release of myxoprincomide-c649 and -c651. Module 11 could be the next rate limiting biosynthetic step. Although functional, it is possible that the module is catalytically inefficient and perhaps as not to “waste” the already assembled peptide chain stalled at module 10, module 11 is either skipped to form majority of the myxoprincomide variants or the growing chain is extended by the upstream module 9 to generate the larger myxoprincomides.

The role of the loading module in terms of structural diversity was not touched on in this work. None of the myxoprincomides presented contained an acyl moiety attached to the N-terminus of the compounds due to the putatively inactive C domain in module 1. By exchanging this inactive C domain with one that has the full set of conserved motifs (e.g. C domains from module 4 and module 8 which incorporate L-serine), it might be possible to observe N-terminal acylated myxoprincomide variants. It is also possible that N-acylated myxoprincomides are produced at the early stage of growth and are

cleaved off later on. A production kinetics study of the overproducer strain and the knockout mutant could shed light into this possibility.

Obtaining sufficient purified myxoprincomides is advantageous for both structural elucidation and bioactivity assays. The possibility of obtaining enough amounts can be achieved by increasing the volume of the VY/2 fed batch cultivation. Whereas it cannot be guaranteed that all myxoprincomides can be successfully isolated, getting a handful pure is already a step in the right direction. For example, it is feasible, using the isolation protocol employed, to isolate myxoprincomide-c651 and myxoprincomide-c562 (data not shown). Unfortunately, the yield was too low for either NMR analysis or bioactivity assays when both compounds were purified. The idea of putting more effort into isolating the myxoprincomide variants comes from the observation that myxoprincomide-c708 imparted a weak activity against *Mucor hiemalis* as opposed to myxoprincomide-c506 which had no observed activity at all. Therefore, it cannot be completely ruled out that the other myxoprincomide variants exhibit antimicrobial activity.

Mxp<sub>1622</sub> and the Group I myxoprincomides fit well under the diversity-oriented paradigm of secondary metabolite biosynthesis wherein a lone hybrid synthetase is capable of generating a suite of structural variants.<sup>[14]</sup> The one pathway – multiple products is common to known *M. xanthus* compounds, where each compound class has at least 2 identified members.<sup>[1]</sup> Chemical diversity from one pathway could be viewed as the organism's efficient strategy for generating a bioactive compound library.<sup>[14]</sup> Alternatively, it could be influenced by the need for enzymatic productivity leading to substrate promiscuity and unconventional biosynthetic steps (i.e. shunt product release or retromodular stuttering).

The structural diversity of the myxoprincomides is further expanded by the Group II myxoprincomides produced by Mxp<sub>897</sub>. It is presumed that the Mxp<sub>897</sub> and Mxp<sub>1622</sub> are mutually exclusive based on the distinct production of Group I and Group II myxoprincomides within the *M. xanthus* species. Whether the diversification of the multimodular assembly lines occurred by vertical inheritance or horizontal gene transfer is unknown and is beyond the scope of this work. However, since there are now a handful of *M. xanthus* genome sequences available in house, the tracking of the genetic evolution of the Mxp megasynthase within the species is feasible.

### 6.3 Althiomycin

The broad spectrum antibiotic althiomycin is produced by the myxobacterial strain *M. xanthus* DK897. Herein, it was shown that the structural backbone of althiomycin is a product of a NRPS-PKS hybrid assembly line composed of AlmA and AlmB. The formation of the mature althiomycin involved three tailoring reactions: a methylation by AlmC, formation of an oxime group by AlmD, and heterocyclization to form the methoxypyrrolinone ring, possibly by AlmF.

Targeted gene inactivation with promoter insertion and subsequent genetic complementation via insertion of the native gene in the *attB* phage attachment site insinuates that AlmC and AlmD act on the structural backbone either during or post assembly. As the peptide intermediate was not detected in the LC-MS profile of the crude extracts after targeted inactivation of either AlmC or AlmD, two possible scenarios were proposed. If the tailoring reactions occur during assembly, it is likely that the downstream domains do not recognize the intermediate without the modifications thereby halting the assembly process. On the other hand, should the tailoring reactions occur post assembly, it is possible that the transporter AlmE does not recognize the peptide intermediate and is therefore not secreted out of the cell.

AlmD warrants further investigation as enzymatic oxime formation has not been mechanistically elucidated. Only the cytochrome P450 NocL of the nocardicin biosynthetic pathway has been ascribed to form oxime functional groups.<sup>[15, 16]</sup> But AlmD distinguishes itself from NocL, as AlmD is not a cytochrome P450 and most likely installs the rarely observed oxime moiety via a different mechanism.

The formation of the methoxypyrrolinone ring by AlmF is still speculative. It is clear that loss of the AlmF functionality led to a significant decrease in althiomycin production. However, the exact role that AlmF plays in the formation of the methoxypyrrolinone ring or in althiomycin biosynthesis altogether is yet to be established.

AlmE showed significant homology to transporters of the major facilitator superfamily. Members of these secondary carriers make use of ion gradients for energy and counts drug efflux as one of its functions.<sup>[17]</sup> Specifically, AlmE showed similarity with

## 5 Conclusion & Outlook

---

bicyclomycin resistance protein which, in *E. coli*, confers resistance to sulfathiazole.<sup>[18]</sup> The common thiazole element between sulfathiazole and althiomycin is most likely necessary for molecular recognition for the drug efflux of these potent antibiotics.

Inactivation of *almE* in the native producer DK897 was unsuccessful, but heterologous expression of *almE* in althiomycin-sensitive *M. xanthus* DK1622 established AlmE's role. Using purified althiomycin and dehydrated althiomycin as test substances, the loss of inhibition zones in the *almE*-harboring DK1622 mutant lawns provided sufficient evidence in AlmE's involvement in self resistance. Inactivation of *almE* in DK897 would have led to a polar effect on the downstream biosynthetic gene cluster as *alm* appears to be a single operon. Assuming there was no polar effect and expression of the althiomycin biosynthetic genes were to proceed, then it could be speculated that the loss of AlmE activity could result in a build up of althiomycin in the cytosol and is potentially lethal to the organism.

More than a year after the *alm* biosynthetic pathway from DK897 was published, an althiomycin pathway homolog was also reported from the enteric bacteria *Serratia marascens* Db10 (gene cluster termed *alb*).<sup>[19]</sup> The gene organization of *alb* matches that of the DK897 *alm* gene cluster. The gene products of *alb1-5* showed significant identity and similarity with *almEDCAB*. Results of deletion experiments on the genes encoding the assembly line (*alb4* and *alb5*) and the tailoring enzymes (*alb2* and *alb3*) parallels the result that were already established for DK897. The heterologous expression of *alb1* in a related althiomycin sensitive strain conferred resistance to the recombinant strain similar to what was observed when *almE* was heterologously expressed in DK1622. While the *almE* gene could not be inactivated in DK897, a deletion mutant was successfully constructed in Db10 which was unable to produce althiomycin owing to the low level of transcription of the other *alb* genes. However, the authors do raise the possibility of second site mutations that could lead to the loss of althiomycin production.

The main difference between *alm* and *alb* lies in the gene products of *almF* and *alb6*. Whereas *almF* is translationally coupled to *almB* and encodes a protein with similarity to proline iminopeptidases, there is a 146 bp gap between *alb5* and *alb6* and *alb6* is predicted to have a thioesterase II which is hypothesized to have a proofreading function. Inactivation of both genes in the respective species did not completely abolish althiomycin biosynthesis but significantly reduced althiomycin production. Initially,



AlmF was hypothesized to be involved in the methoxypyrrolinone ring. Supposing that the AlmB TE domain is capable of ring formation by itself, an alternative function for AlmF would be the regeneration of PCP domains during biosynthesis if the PCP is tethering an incorrect intermediate (e.g. an *R* configuration at C4 of the thiazoline ring instead of an *S* configuration). If AlmF does have a proline iminopeptidase-like activity (exclusive cleaving of L-proline residues from peptides), then AlmF could recognize the thiazoline ring with an *R* configuration at C4 based on its similarity to the heterocyclic structure of proline and consequently hydrolyze it from the thioester. An AlmF enzymatic assay using althiomycin and epialthiomycin (althiomycin with an *R* configuration at C4 of the thiazoline ring)<sup>[20]</sup> or their corresponding intermediates could shed light if proofreading of the thiazoline group is indeed the function of AlmF.

Extrapolating from the secondary metabolome survey of *M. xanthus* strains, althiomycin appear to be an uncommon antibiotic produced by this species.<sup>[1]</sup> It is possible that the strain DK897 acquired the *alm* gene cluster by horizontal gene transfer as a consequence of the microbial predatory nature of myxobacteria.<sup>[21]</sup>

## 6.4 Final word

Genetically tractable and relatively easy to cultivate, *Myxococcus xanthus* has now presented itself, not just as an excellent model for social behavior in myxobacteria but as a representative strain for the actualization of bacterial secondary metabolite biosynthetic potential. With continuous efforts, it is possible to obtain the full secondary metabolome by applying and extending the techniques presented in this work.

## 6.5 References

- [1] D. Krug, G. Zurek, O. Revermann, M. Vos, G. J. Velicer, R. Müller, *Applied and Environmental Microbiology* **2008**, 74, 3058-3068.
- [2] D. Krug, G. Zurek, B. Schneider, R. Garcia, R. Müller, *Analytica Chimica Acta* **2008**, 624, 97-106.
- [3] C. Volz, C. Kegler, R. Muller, *Chem Biol* **2012**, 19, 1447-1459.
- [4] B. R. Villiers, F. Hollfelder, *ChemBioChem* **2009**, 10, 671-682.

- [5] N. Fusetani, S. Matsunaga, H. Matsumoto, Y. Takebayashi, *Journal of the American Chemical Society* **1990**, 112, 7053-7054.
- [6] F. Itagaki, H. Shigemori, M. Ishibashi, T. Nakamura, T. Sasaki, J. Kobayashi, *Journal of Organic Chemistry* **1992**, 57, 5540-5542.
- [7] L. Chill, Y. Kashman, M. Schleyer, *Tetrahedron* **1997**, 53, 16147-16152.
- [8] E. W. Schmidt, C. Raventos-Suarez, M. Bifano, A. T. Menendez, C. R. Fairchild, D. J. Faulkner, *Journal of Natural Products* **2004**, 67, 475-478.
- [9] N. Geib, K. Woithe, K. Zerbe, D. B. Li, J. A. Robinson, *Bioorganic & Medicinal Chemistry Letters* **2008**, 18, 3081-3084.
- [10] J. Moldenhauer, X. H. Chen, R. Borriss, J. Piel, *Angew Chem Int Ed Engl* **2007**, 46, 8195-8197.
- [11] J. Cortes, K. E. Wiesmann, G. A. Roberts, M. J. Brown, J. Staunton, P. F. Leadlay, *Science* **1995**, 268, 1487-1489.
- [12] L. Li, W. Deng, J. Song, W. Ding, Q. F. Zhao, C. Peng, W. W. Song, G. L. Tang, W. Liu, *Journal of Bacteriology* **2008**, 190, 251-263.
- [13] M. Juguët, S. Lautru, F. X. Francou, S. Nezbedova, P. Leblond, M. Gondry, J. L. Pernodet, *Chemistry & Biology* **2009**, 16, 421-431.
- [14] M. A. Fischbach, J. Clardy, *Nature Chemical Biology* **2007**, 3, 353-355.
- [15] M. Gunsior, S. D. Breazeale, A. J. Lind, J. Ravel, J. W. Janc, C. A. Townsend, *Chemistry & Biology* **2004**, 11, 927-938.
- [16] W. L. Kelly, C. A. Townsend, *Journal of Bacteriology* **2005**, 187, 739-746.
- [17] S. S. Pao, I. T. Paulsen, M. H. Saier, Jr., *Microbiology and Molecular Biology Reviews* **1998**, 62, 1-34.
- [18] B. P. Nichols, G. G. Guay, *Antimicrobial Agents and Chemotherapy* **1989**, 33, 2042-2048.
- [19] A. J. Gerc, L. Song, G. L. Challis, N. R. Stanley-Wall, S. J. Coulthurst, *PLoS One* **2012**, 7, e44673.
- [20] K. Inami, Shiba, T., *Bulletin of the Chemical Society of Japan* **1986**, 59, 2185-2189.
- [21] G. J. Velicer, M. Vos, *Annual Review of Microbiology* **2009**, 63, 599-623.

## 7 Author's Contribution to Chapters 3 to 5

### Chapter 3:

The author carried out the mutagenesis, strain cultivation, sample work up, HPLC-MS data and statistical analyses. Further the author performed the optimization of myxoprincomide-c506 production inclusive of selection of best producer, genetic manipulation, and improvement of cultivation conditions. The author performed the sequence analysis on the presented biosynthetic gene clusters and designed and carried out gene inactivation experiments of possibly related genes. The author was involved in the purification of myxoprincomide-c506 and performed the feeding experiments.

### Chapter 4:

The author conceptualized and carried out the optimization of cultivation conditions to improve access to the newly identified myxoprincomides. The author purified myxoprincomide-c534 and myxoprincomide-708 and performed the structural elucidation from tandem MS and NMR data. The author combined feeding of isotope-labeled substrates and tandem MS experiments for the structure elucidation of the less abundant myxoprincomides. The author performed the genome scan for related genes and carried out the inactivation experiments.

### Chapter 5:

The author performed the genome scanning for the althiomycin biosynthetic gene cluster in the strain *M. xanthus* DK897 and the corresponding gene inactivation experiments. The author carried out the heterologous expression of *almE* in *M. xanthus* DK1622 and performed the althiomycin sensitivity assay.